

WORLD INTELLECTUAL PROPERTY ORGANIZATION



INTERNATIONAL APPLICATION PUBLY HED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5: C12Q 1/00, 1/48, C12N-9/99 G01N 33/68, C07K 13/00 A61K 37/64, C07K 15/00

(11) International Publication Number:

WO 93/02209

(43) International Publication Date:

4 February 1993 (04.02.93)

(21) International Application Number:

PCT/US92/06137

A1

(22) International Filing Date:

22 July 1992 (22.07.92)

(74) Agents: ALTMAN, Daniel, E. et al.; Knobbe, Martens, Olson & Bear, 620 Newport Center Drive, Suite 1600,

Newport Beach, CA 92660 (US).

(30) Priority data:

735,614

22 July 1991 (22.07.91)

US

(71) Applicant: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA [US/US]; 300 Lakeside Drive, 22nd Floor, Oakland, CA 94612-3550 (US).

(72) Inventors: SOWADSKI, Janusz, M.; 7025 Charmant Drive, Apt. 264, San Diego, CA 92122 (US). TAYLOR, Susan, S.; 1408 Via Alta, Del Mar, CA 92014 (US). KNIGHTON, Daniel, R.; 3920 Carmel Springs Way, San Diego, CA 92130-2254 (US).

(81) Designated States: AT, AU, BB, BG, BR, CA, CH, CS, DE, DK, ES, FI, GB, HU, JP, KP, KR, LK, LU, MG, MN, MW, NL, NO, PL, RO, RU, SD, SE, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LU, MC, NL, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, SN, TD, TG).

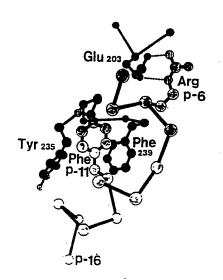
Published

With international search report.

(54) Title: METHODS OF DESIGNING SPECIFIC AFFECTORS USING THREE-DIMENSIONAL CONFORMATION OF ENZYME/AFFECTOR COMPLEX

(57) Abstract

The present invention includes methods for rational drug design. One exemplary method disclosed herein teaches the preparation of a highly specific affector of a first enzyme when the first enzyme is a member of a class of enzymes having a conserved catalytic core with a plurality of invariant residues. This method comprises identifying a second enzyme that is a member of that class of enzymes and has a known affector. The affector can be an inhibitor or activator of the second enzyme. In the practise of the method, a first complex is formed between the second enzyme and the known affector and data is obtained regarding the three-dimensional coordinates of the invariant residues in the complex. These coordinates are used to form a template. A model is then generated in which the first enzyme is in a conformation with the invariant residues in substantially the same conformation as in the template. Changes in the variable residues of the catalytic core of the first enzyme are compared to the variable residues in the catalytic core of the second enzyme. The second enzyme is modified to include these non-conserved changes and an affector is designed using computer modelling that will induce a conformation on the catalytic core of the first enzyme in which the invariant residues are in substantially the same coordinates as the coordinates of the template when the first enzyme is formed as a second complex with the newly designed affector. The designed affector can be further refined to provide improved affector activity.





FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FI	I-inland	MI.	Mali
ΑU	Australia	FR	France	MN	Mongolia
88	Barbados	GA	Ciabon	MR	Mauritania
BE	Belgium	GB	United Kingdom	MW	Malawi
8F	Burkina Faso	GN	Guinea	NL	Netherlands
BG	Bulgaria	GR	Greece	NO	Norway
BJ	Benin	HU	Hungary	PL	Poland
BR	Brazil	IE	Ireland	RO	Romania
CA	Canada	iT	Italy	RU	Russian Federation
CF	Central African Republic	JP	Japan	SD	Sudan
CG	Congo	KP	Democratic People's Republic	SE	Sweden
CH	Switzerland		of Korca	SN	Senegal
Cl	Côte d'Ivoire	KR	Republic of Korea	SU	Soviet Union
CM	Cameroon	LI	Licchtenstein	TD	Chad
C?	Czechoslovakia	LK	Sri Lanka	TG	Togo
DE	Germany	LU	Luxembourg	US	United States of America
DK	Denmark	MC	Monaco		
ES	Spain	MG	Madagascar		

METHODS OF DESIGNING SPECIFIC AFFECTORS USING THREE DIMENSIONAL CONFORMATION OF ENZYME/AFFECTOR COMPLEX

Field of the Invention

The present invention relates to rational design of specific affectors for a given enzyme using data obtained regarding the three dimensional conformation of an enzyme/affector complex. More particularly, it relates to such methods wherein the conformation of the conserved catalytic core of a given enzyme class is elucidated and highly specific affector molecules for a particular member of that class are designed.

Background of the Invention

10

5

1

Drug design based on an analysis of the structural features of a molecule is still in its infancy. At present, an analysis of X-ray crystallographic data at best permits the design of broadly acting affector molecules. While these affector molecules can be further refined to impart some selectivity, affector design does not produce molecules having the fine tuned specificity of, for example, an antibody for its antigen. This level of selectivity control is not always necessary; however, therapeutic regimes directed to the control of enzymes involved in certain cancers, genetic disorders, and infectious agents will require this type of selectivity.

20

15

Enzymes can be classified into broad families or classes having similar activities, with each enzyme having a specific function. For example, many proteins phosphorylate their substrate. These enzymes are broadly labelled as kinases. A myriad of kinases exist for a myriad of functions. Within this broad group, kinases can be subgrouped based on similarities in substrate, requirements for additional cofactors or similar amino acid residues that are targets for phosphorylation.

25

Within any given cell, there may be many active members of a given enzyme family. If one member of the family shows aberrant activity, then it may be therapeutically advantageous to alter the activity of this single enzyme to the exclusion of other similar or related enzymes. Such is the case for the protein kinase family where aberrant phosphorylation events can be associated with abnormal cell growth and regulation. This is observed in proto-oncogene related cancers. For example, the pp60^{c-src} protein, needs to be controlled to the exclusion of other protein kinases in order to maintain normal cell metabolism.

30

Protein phosphorylation as a mechanism for regulating protein activity was first recognized in 1955 with glycogen phosphorylase. Protein phosphorylation and dephosphorylation is widespread and impacts nearly all aspects of growth and homeostasis

5

10

15

20

25

30

7

in the eukaryotic cell. Protein kinases catalyze the transfer of the γ -phosphate of MgATP to a protein substrate. The protein kinases, constitute a large and very diverse family of enzymes. Although these enzymes differ in size, substrate specificity, mechanism of activation, subunit composition, and subcellular localization, all, nevertheless, share a homologous catalytic core that has been conserved throughout evolution.

It is not yet possible to regulate a given enzyme at will. While there are hundreds of different protein kinases, only a few of these can be readily purified. Moreover, even among those enzymes that can be purified, many cannot be used for X-ray crystallographic studies. The sequences of many enzymes have been cloned and expressed; however, not all of these are chemically active. Therefore, even if a molecule that cannot be readily purified is cloned and expressed, it may not be functional and thus, would not provide an adequate model for structural studies. Further, even if a recombinant protein is functional, it may not be readily crystallizable. These and other roadblocks have heretofore prevented the design or identification of affector molecules directed to a particular enzyme. Thus, heretofore it has not been possible to provide a method for the design of affector molecules for a given member of an enzyme family.

Brief Description of the Figures

Figure 1 illustrates the general topological arrangement of the catalytic subunits from a number of different protein kinases.

Figure 2 diagrams the placement of the catalytic region within various members of the protein kinase family.

Figure 3 is a stereo view of the electron density for the structure determination. Figure 3A provides the density calculated to 2.7 Å. Figure 3B provides the density calculated with 10.0 to 2.7 Å refined model phases.

Figure 4 is a stereo view of the C- α backbone and includes twenty residues of PKI(5-24).

Figure 5 provides data on the location and orientation of MgATP. Figure 5A illustrates the general localization of MgATP. Figure 5B is a close-up of difference density enclosing an oriented model of AMP.

Figure 6 is an overall two dimensional topology diagram for the C-subunit. of cAPK. Figure 7 provides stereo views of selected conserved areas.

Figure 8 illustrates the conserved catalytic core of c-AMP dependent protein kinase. Figure 8A is a space-filling model of the catalytic core. Figure 8B is a diagram of the

conserved catalytic core using the RIBBON program of the PAP package. Figure 8C is a space-filling model identical to A, but includes PKI(5-24).

Figure 9 diagrams the conformation of bound PKI (5-24).

5

10

15

20

25

30

Figure 10 illustrates the high affinity binding site interactions between the catalytic subunit and the inhibitor peptide. In Figure 10A the P-11 Phe ring (numbered 356) is shown in relation to three nearby hydrophobic residues of the C-subunit. Figure 10B illustrates the interactions of the P-11 Phe and P-6 Arg sidechains with the C-subunit.

Figure 11 summarizes the interactions that contribute to recognition of the common consensus region of the inhibitor peptide.

Figure 12 illustrates the consensus recognition site binding interactions. Figure 12A is an illustration of the electron density corresponding to the anionic P-3 site. Figure 12B illustrates the electron density of the P-2 Arg side chain. Figure 12C illustrates the electron density of the P+1 Ile sidechain.

Figure 13 illustrates the catalytic site area. Figure 13A provides the site of catalysis together with the probable catalytic base sidechain of Asp 166 near the β-C of the P Ala. Figure 13B diagrams the consensus recognition site residues Arg-Arg-Asn-Ala-Ile. Figure 13C shows the phosphate of Thr 197 buried among sidechains of cationic and H-bonding residues.

Figure 14 is a schematic illustrating the relationship of invariant amino acids at the active site.

Figure 15 is a schematic of the conserved and variable regions of the catalytic subunit.

Figure 16 illustrates the amino acids present in PKI(5-24) that provide important interactions with cAPK.

Figure 17 provides a list of the coordinates that define the three-dimensional template.

Figure 18 provides photographs of the crystal forms.

Summary of the Invention

In accordance with one aspect of the present invention, there is provided a method of designing a highly specific affector which exerts an effect on the activity of a first enzyme. In this method, the first enzyme is a member of a class of enzymes having a conserved catalytic core. The method comprises the following steps: identifying a second enzyme that is a member of the class in which a first affector can affect the activity of the second enzyme, forming a first complex of the first affector and the second enzyme, obtaining data

5

10

15

20

25

30

-4-

regarding the conformation of the second enzyme at sites greater than 5 Å from the site of catalysis of the second enzyme in the first complex, and designing an affector which induces a conformation on the first enzyme at sites greater than 5 Å from the site of catalysis thereof which is homologous to the conformation of the second enzyme at homologous sites in the first complex, when the affector is formed as a second complex with the first enzyme. Preferably, this method additionally comprises crystallizing the first complex and obtaining X-ray crystallography data therefrom. In a preferred form of this method, all of the members of the class have related functions, and the catalytic cores of all of the members of the class have conserved amino acid residues. In this form of the method, preferably the designing step comprises designing an affector having homologous topography and charge fields that complement the catalytic core of the first enzyme and capable of inducing a conformation wherein the conserved amino acid residues of the first enzyme are in homologous locations to the second enzyme in the first complex. The affectors can be inhibitors, activators or other affectors of enzyme activity. The first affector can be all or a portion of the first enzyme, and the first complex can be a holoenzyme. The class of enzymes can comprise protein kinases or any other suitable class. The second enzyme can be a viral oncogene product or a cellular homologue thereof, such as p60 v-Src from RSV or its cellular homologue, pp60 c-src. The second enzyme can also be cAMP-dependent protein kinase. The second enzyme can be a native mammalian protein kinase or a recombinant protein kinase. In a preferred form of the method, the designing step comprises identifying a potential affector likely to induce a conformation on the catalytic core of the first enzyme which is homologous to the conformation of the catalytic core of the second enzyme in the first complex, when the affector is formed as a second complex with the first enzyme, and determining whether the potential affector induces the conformation through the use of a technique selected from the group consisting of computer modelling, small angle neutron scattering, and circular dichroism. In this preferred method, the potential affector comprises a peptide, and the potential affector comprises a molecule selected from the group consisting of nucleotides, polynucleotides, alcohols, polymers, lipids carbohydrates, sugars, native or synthetic molecules, protein and organic compounds and combinations thereof. In accordance with this aspect of the invention, the method can include producing the affector. Thus, the present invention also includes the affector produced from the method.

In another aspect of the present invention, there is provided another method of designing a highly specific affector which exerts an effect on the activity of a first enzyme.

5

10

15

20

25

30

-5-

In this method, the first enzyme is also a member of a class of enzymes having conserved residues at an affector binding site. This method comprises the following steps: identifying a second enzyme that is a member of the class in which a first affector can affect the activity of the second enzyme, the first affector having a dissociation constant with the second enzyme of less than 1 µM, forming a first complex of the first affector and the second enzyme, obtaining data regarding the conformation of the affector binding site of the second enzyme in the first complex, and designing an affector which induces a conformation on the affector binding site of the first enzyme which is homologous to the conformation of the affector binding site of the second enzyme in the first complex, when the affector is formed as a second complex with the first enzyme. In one form of this method, the class of enzymes has a nucleotide binding site and each of the affectors is capable of binding to the nucleotide binding site.

In still another aspect of the present invention, there is provided another method of designing a highly specific affector which exerts an effect on the activity of a first enzyme. In this method, the first enzyme is also a member of a class of enzymes having a conserved catalytic core. This method comprises the following steps: identifying a second enzyme that is a member of the class in which a first affector can affect the activity of the second enzyme, forming a first complex of the first affector and the second enzyme, the first complex having at least three points of contact between the first affector and second enzyme, obtaining data regarding the conformation of the catalytic core of the second enzyme in the first complex, and designing an affector which induces a conformation on the catalytic core of the first enzyme which is homologous to the conformation of the catalytic core of the second enzyme in the first complex, when the affector is formed as a second complex with the first enzyme.

Still another aspect of the present invention provides a crystallized protein kinase/affector complex having stable decay characteristics over 15 minutes and a crystallized protein kinase/affector complex having a Bragg spacing diffraction limit of less than 4Å. Preferably, the crystallized protein kinase of this aspect of the invention exhibits both of these characteristics. The present invention also provides a crystallized complex of a cAMP-dependent protein kinase and an inhibitor thereof. This crystallized complex can be used in an X-ray crystallography procedure to produce data regarding the three dimensional structure of the cAMP-dependent protein kinase in the complex, and this data can be used for designing an inhibitor of a second protein kinase which will induce a similar three dimensional structure on the active site of the second protein kinase as the three

5

10

15

20

25

30

-6-

dimensional structure of the cAMP-dependent protein kinase in the complex. Thus, the present invention also includes an inhibitor designed by this method.

Another preferred method of the present invention involves preparing a highly specific affector of a first enzyme, with the first enzyme being a member of a class of enzymes having a conserved catalytic core with a plurality of invariant residues. This method comprises the following steps: identifying a second enzyme that is a member of the class and having a known affector thereof, forming a first complex of the second enzyme and the known affector, obtaining data regarding the three dimensional coordinates of the invariant residues in the first complex, the coordinates forming a template, generating a model wherein the first enzyme is in a conformation in which the invariant residues are in substantially the same conformation as in the template, identifying a change in the variable residues in the catalytic core of the first enzyme in the conformation of the template when compared to the variable residues in the catalytic core of the second enzyme in the conformation of the template, preparing a modified form of the second enzyme, wherein the modified second enzyme includes a non-conserved change identified through this method, and designing an affector which computer modelling indicates will induce a conformation on the catalytic core of the first enzyme in which the invariant residues are in substantially the same coordinates as the coordinates of the template, when the first enzyme is formed as a second complex with the affector designed in this step. Preferably the identified change is a non-conserved change in the variable residues. In a preferred form of this method, the method also includes forming a third complex of the modified second enzyme and an affector capable of binding thereto, obtaining data regarding the three dimensional coordinates of the invariant residues in the third complex, and using the data obtained in the previous step to design an affector which computer modelling indicates will induce a conformation on the catalytic core of the first enzyme in which the invariant residues are closer to the coordinates of the template than the conformation induced by the affector designed previously, when the first enzyme is formed as a fourth complex with the affector designed in this step. The affector used for computer modelling can be the known affector. Preferably, the method also includes modifying the computer modelling in light of the data obtained through the method prior to designing the affector. Amino acid sequence data relating to the catalytic cores of the first and second enzymes is preferably obtained. Site directed mutagenesis of a recombinantly produced second enzyme can be used in accordance with the method. In one preferred aspect of this method, the coordinates of the template are substantially as shown in Figure 17 and the template can include

-7-

coordinates separated by the distances substantially as shown in Table 4. The affectors can be inhibitors or other affectors. The method can also include preparing the designed affector. Thus, the present invention also includes the affectors prepared through this method, and also includes pharmaceutical compositions containing these affectors.

5

The present invention also includes a method of designing a specific inhibitor for a protein kinase, comprising the following steps: obtaining data regarding the three-dimensional structure of a first protein kinase, and using the data in the design of an inhibitor for a second, different, protein kinase. The first protein kinase is preferably cAMP dependent protein kinase or an analogue thereof. The obtaining step preferably comprises crystallizing a complex of cAMP dependent protein kinase and an inhibitor thereof, and additionally includes obtaining X-ray crystallography data on the crystallized complex obtained from the crystallizing step. Thus, information regarding the structure of the crystallized complex obtained in the crystallizing step through circular dichroism, small angle neutron scattering, or other chemical procedures can be obtained.

15

10

In a preferred form of the present invention, there is provided the use of the data of Figure 17 or of Table 4 in the design of an affector for a protein kinase.

20

25

30

Still another aspect of the present invention involves a method of preparing a highly specific inhibitor of a first enzyme. The first enzyme is a member of a class of enzymes having a conserved catalytic core with a plurality of invariant residues. This method includes the following steps: (a) identifying a second enzyme that is a member of the class and having a known first inhibitor thereof, (b) forming a first complex of the second enzyme and the first inhibitor, (c) obtaining data regarding the three dimensional coordinates of the invariant residues in the first complex, (d) designing a second inhibitor which computer modelling indicates will induce a conformation on the catalytic core of the first enzyme in which the invariant residues are in substantially the same coordinates as the coordinates obtained in step (c), when the first enzyme is formed as a second complex with the second inhibitor, (e) preparing the second inhibitor, (f) forming a third complex of the second inhibitor and a third enzyme complexable therewith, the third enzyme having a plurality of the invariant residues, (g) obtaining data regarding the three dimensional coordinates of the invariant residues in the third complex, and (h) using the data obtained from step (g) to design a third inhibitor which computer modelling indicates will induce a conformation on the catalytic core of the first enzyme closer to that in which the invariant residues are in substantially the same coordinates as the coordinates obtained in step (c) than induced by the second inhibitor, when the first enzyme is formed as a fourth complex with the third

5

10

15

20

25

30

inhibitor. This first inhibitor is in one embodiment of this method an inhibitory domain of the second enzyme. The third enzyme preferably contains at least 5 invariant residues, and can be a naturally occurring enzyme or a mutant enzyme.

Still another aspect of the present invention involves a method of determining the efficacy of a first affector in affecting a first enzyme that is a member of a class of enzymes having a plurality of invariant residues among the members of the class. This method includes the following steps: determining the three dimensional coordinates of the invariant residues of a second enzyme in a first conformation wherein the second enzyme is in a complex with a second affector that is a strong affector of the enzyme, determining the three dimensional coordinates of the invariant residues of the second enzyme in a second conformation wherein the enzyme is in a conformation other than the first conformation, identifying the mobile invariant residues of the enzyme, the mobile invariant residues being those invariant residues at coordinates substantially different in the first conformation than in the second conformation, determining the three dimensional coordinates of the mobile invariant residues of the first enzyme when the first enzyme is in a conformation wherein the first enzyme is in a complex with the first affector, comparing the three dimensional coordinates of the mobile invariant residues of the first enzyme in the conformation with the coordinates of the mobile invariant residues of the enzyme in the first conformation, wherein relatively small differences in the coordinates indicate relatively high efficacy of the first affector. The step of determining the coordinates of the first enzyme in the conformation is preferably performed using computer modelling of the conformation. The steps of determining the first and second conformations preferably comprise obtaining X-ray crystallographic data of the enzyme. The second conformation can be a conformation produced by a ternary complex, such as one comprising a protein kinase, a nucleotide and an affector. The second conformation can also be a conformation produced by the second enzyme not complexed with a ligand, or the same enzyme as the first enzyme.

In an additional aspect of the present invention, there is provided a method of designing specific inhibitors of a first protein kinase having a template defined by a plurality of invariant residues present in substantially all protein kinases. This method comprises the following steps: obtaining data regarding the three dimensional coordinates of the invariant residues of a second protein kinase and of the points of contact between the second protein kinase and a known inhibitor thereof, the coordinates being obtained when the second protein kinase is formed as a complex with the known inhibitor, generating a model of the first protein complex wherein the template is defined by the positions of the invariant

5

10

15

20

25

30

residues in the complex, examining the amino acid residues present in the first protein kinase at positions corresponding to the points of contact in the complex, and designing an inhibitor of the first protein kinase capable of forming ionic and hydrophobic interactions with the amino acid residues. The method of Claim 62, wherein the second protein kinase is cAMP dependent protein kinase. The known inhibitor can be PKI(5-24). For this known inhibitor, the points of contact in the complex preferably comprise positions corresponding to the positions selected from the group consisting of P+1, P-2, P-3, P-6 and P-11 along the known inhibitor. The positions corresponding to the points of contact in the examining step preferably comprise positions within a sphere having a radius of 11 Å, more preferably 6 Å, from the coordinates of the point of contact obtained in the obtaining step. The designing step preferably additionally comprises designing the inhibitor to form appropriate hydrogen bonding with the amino acid residues.

Further details concerning the present invention are provided in the following detailed description.

Detailed Description of the Invention

CITED REFERENCES INCORPORATED BY REFERENCE

A number of articles are specifically cited herein as providing background information useful, but not essential, to those of ordinary skill in the art in the practice of the present invention. As such, the disclosure of each of these articles is hereby explicitly incorporated by reference.

INTRODUCTION

The protein kinase family of enzymes is used as a model for this invention. These enzymes are involved at all levels of regulation in the eukaryotic cell. They act as "transistors" for the cell, receiving signals and amplifying the message inside the cell. Protein kinases receive hormone signals from outside the cell. They are involved in cell growth, for cellular homeostasis, and for triggering the steps of mitosis.

In addition, many oncogenes code for protein kinases. These oncogenic protein kinases are also very diverse in their structure and location within the cell. However, all are derived from normal cellular components and all, in one way or another are defective in their ability to be turned off. In other words, they are constitutively active in contrast to their protooncogene counterparts which are turned off in the absence of the appropriate signal. Thus, protein kinases are not only an essential part of normal cell growth and division, but, can lead to oncogenesis when their normal function becomes genetically impaired.

5

10

15

20

25

30

Diversity is a hallmark of the protein kinase family. For example, growth factor receptors, such as the insulin receptor, are large proteins with a major extracellular domain for binding growth factor, a single membrane spanning domain, and an intracellular protein kinase domain that is activated in response to growth factor binding. The kinase activity is limited to a specific domain of the protein. Control of the insulin receptor may play an important role in the control of diabetes. Protein kinase C is activated by diacyl glycerol and Ca²⁺ and is also activated by the tumor promoting phorbol esters. It is a cytoplasmic protein that in its active state is associated with the plasma membrane. Another protein kinase, cdc2, associates with cyclin B and is an essential trigger for mitosis. The transforming protein in Rous Sarcoma Virus, pp60^{v-src} is anchored to the cytoplasm surface of membranes. In spite of the diversity in size, subunit composition, location in the cell, and mechanism of activation, all protein kinases share a common enzymatic activity and a conserved catalytic core, indicating that all have likely evolved from a common functional precursor. Thus, one aspect of the present invention provides a method for developing highly selective inhibitors for members of the protein kinase family.

The first protein kinase to be purified was phosphorylase kinase. The second was phosphorylase kinase kinase, later renamed cAMP-dependent protein kinase (EC2.7.1.37:ATP:protein serine phosphotransferase) when its broader substrate specificity was appreciated. Not only was cAMP-dependent protein kinase (cAPK) one of the first protein kinases to be characterized, it also is one of the simplest and best understood biochemically. Its simplicity is due primarily to its mechanism of activation, which involves subunit dissociation. With the exception of the oncogenic enzymes, all protein kinases typically are maintained in an inactive state in the absence of the appropriate activating signal. In the case of cAPK, the ligand triggering activation is cAMP, one of the first recognized second messengers for hormone signalling. In the absence of cAMP, the enzyme is sequestered as an inactive holoenzyme containing two regulatory (R) and two catalytic (C) subunits. When intracellular cAMP levels are elevated, the cyclic nucleotide binds to the R-subunit, thus causing the complex to dissociate into a R2 dimer and two free and active C-subunits. The general consensus sequence recognized by the C-subunit is Arg-Arg-X-Ser/Thr-Y, where X is any small residue and Y is a large hydrophobic group. The conserved catalytic core found in all protein kinases is contained within this relatively simple monomeric C-subunit.

This invention provides the first crystal structure of a protein kinase with its catalytic subunit intact. Knowledge of the conformation of the catalytic structure of cAPK is central

to the understanding of protein kinase activity. Not only is the structure of the cAMP-dependent protein kinase catalytic site provided, but, the crystals contain a bound inhibitor peptide. This inhibitor peptide, PKI(5-24), is a fragment of the heat stable protein kinase inhibitor (PKI). This peptide includes the consensus features common to all peptide substrates and inhibitors of cAMP-dependent protein kinase. In addition, it contains other features that convey unique high affinity binding characteristics. Thus, precise properties of binding and interaction are described. From this data, a template is derived from which all other protein kinases can be modelled and from which other inhibitors can be designed.

5

10

15

20

25

30

One of the more important questions regarding protein phosphorylation is how the targeted protein substrate is recognized by a specific protein kinase. This question has remained particularly elusive until now because the determinants for peptide recognition are widely dispersed and in some cases well-removed from the actual site of phosphotransfer. Owing to its simplicity as well as its relative ease of purification, the catalytic or C-subunit of cAMP-dependent protein kinase serves here as a prototype for identifying functional sites that are involved in substrate recognition and catalysis. Chemical analyses and procedures, such as affinity labeling, group specific labeling, and fluorescence energy transfer all have provided clues about regions involved in peptide recognition, MgATP binding, and catalysis. Substrate analogues provide indirect information about binding sites important for affector molecule specificity. Further NMR, circular dichroism, small angle neutron scattering (SANS) and other chemical procedures offer further insight into the structure of the enzyme. However, X-ray crystallography provides a comprehensive three dimensional structure that can confirm and integrate these other techniques.

The expression of the C-subunit in E. coli not only facilitated these structural studies, but also has permitted recombinant approaches to be used to further modify the active site of cAPK and thereby mimic the reactive site of other protein kinases. Information to aid in these studies is obtained from sequence data available for the protein kinase family. Hanks et al., Science 241: 42, 1988, is one source of such data. Such sequence comparisons have identified highly conserved regions including several invariant residues, variable regions, and places where inserts and deletions can be tolerated. Both chemical and sequence information are used here to verify the structure data obtained from the X-ray diffraction studies. As will be disclosed herein, this body of information permits the design of other affector molecules specific for other protein kinases. Further, this information serves as guidelines for the design of specific affector molecules for enzymes from a wide variety of enzyme families.

The existing basis for the design of specific inhibitors for protein kinases, in the absence of the three dimensional structure provided herein, relies on the use of synthetic peptides based primarily on the sequences of known substrates and inhibitors. In the case of cAMP-dependent protein kinase, there are some very specific high affinity peptides available. Existing inhibitors also include nucleotide and nucleoside derived compounds found through traditional means. However, these nucleoside and nucleotide inhibitors do not generally exhibit the type of specificity observed with peptide inhibitors. In general, such specific peptide inhibitors are not available for other protein kinases. Specificity for cAMP-dependent kinase improves with the addition of amino acids postulated to lie outside of the catalytic core. We have discovered that these regions are also important for inhibitor design. Knowledge of these sites provides a "lock" to permit for the first time the tailoring of inhibitors for any given protein kinase. Thus, one important aspect of the invention lies in the design of the "lock", that requires an understanding of the three dimensional structure of the complex of the catalytic subunit of cAMP-dependent protein kinase, with its very potent specific inhibitor, PKI(5-24).

Disclosed herein is a template gleaned from the crystal structure of the catalytic subunit of cAMP-dependent protein kinase. Just as the chemical information derived from the C-subunit serves as a framework for interpreting the entire kinase family, the structure of cAPK provides information for the creation of a template for viewing the conserved catalytic core of all eukaryotic protein kinases. This invention further provides a model for the identification and design of molecules capable of interacting with the catalytic core of a given enzyme by analyzing the conserved catalytic core of another member of that enzyme class.

X-RAY CRYSTALLOGRAPHY

5

10

15

20

25

30

X-ray crystallography permits three dimensional molecular analysis of a protein at the atomic level. Analysis requires the production of crystals and crystal production requires a pure concentrated product. Further, complexes of a protein of interest together with a second interacting molecule provides information on the conformational changes occurring within a protein in response to that second molecule. X-ray crystallography of a protein with its substrate, an antibody or a drug can provide information for rational drug design.

An X-ray diffraction pattern taken from a crystal looks like an array of spots of varying intensities. Each spot is related to one of the Fourier coefficients of the electron density pattern in the crystal. Thus, the electron density in the crystal can be reconstructed if a sufficient number of diffraction spots can be measured and the relative phase angles of

5

10

15

20

25

30

the Fourier coefficients can be determined. Thus, a crystallized enzyme used in the practice of certain aspects of the present invention should be of sufficient quality to obtain these measurements. For example, the spots of varying intensity in the diffraction pattern decay over time. It is quite difficult to work with diffraction patterns with half lives of less than 10 hours. However, it is possible to work with diffraction patterns having half lives as short as about 15 minutes to 3 hours, depending on the amount of structural data desired to be obtained. Further, it is believed possible to work with crystals of even shorter half lives using equipment and computer programs more advanced than commonly available today. Additionally, not all crystals are of equal quality and poor crystals have large Bragg spacing diffraction limits. Thus, a workable crystal should have a Bragg spacing diffraction limit of less than 4 Å.

Determination of phase angles uses isomorphic replacement to insert atoms into defined positions in the crystal for diffraction data measurement. These angles provide information that permit the production of an electron density map. The map is then used to build an atomic model from which three-dimensional coordinates are measured that define the structure of the crystallized molecule.

MODEL SYSTEM

X-ray crystallography has been employed for the rational design of drugs and other interacting molecules. However, to date, the rational design of affector molecules has been limited to a study of the active site of the protein/affector molecule interaction. Potential affectors designed from this information have not been obtained by looking at interactions beyond the active site. We believe that these interactions assist in binding and thereby contribute to binding specificity. Thus, we have discovered that it is these interactions in concert with information obtained from the active site that make the design of specific affector molecules a possibility. Moreover, this information additionally permits the design of specific affector molecules for related but nonidentical enzymes.

Many enzymes within a cell have evolved from common progenitors. These enzymes share common enzymatic activities and one example is the protein kinase family. Since the functions of the enzymatic families or classes are broadly conserved, at least a portion of the catalytic site is also conserved. Therefore to a large extent rational drug design relies on the identification of the familial similarities and hence drugs are designed to react broadly within a given family or class.

While all members of an enzymatic class may provide a similar activity, such as phosphorylation or dephosphorylation, each member may have only one specific target.

5

10

15

20

25

30

Thus, successful rational drug design based on group similarities would provide molecules that also interact broadly. Where there are many members of an enzymatic class acting specifically within a restricted locale such as a single cell, a broadly acting drug would interact with any number of enzymes from the group. Thus, the interaction would be general and not specific. In contrast, the present invention advantageously is capable of providing affectors with highly specific interactions for a given member of an enzyme class.

Previous methods for rational drug design require the crystallization of the target molecule of interest. However, the production of useful crystals is both difficult and time consuming. It first depends on the ability of the target molecule to be isolated and purified in sufficient quantity for crystallization. A large number of crystallization conditions often need to be tested and once a crystal is made that is of sufficient quality, additional crystals often need to be produced in order to have enough material for analysis. Further, not all molecules are readily purified or readily crystallized. Advantageously, the present invention discloses a method whereby only one enzyme within a family of enzymes need be crystallized.

This invention teaches a method for the identification and design of specific molecules interacting with a specific enzyme wherein the specific enzyme is a member of a broadly acting enzymatic group or class.

The particular enzyme class chosen for this invention is preferably one that has the characteristics generally associated with an enzyme class developed from a divergent evolutionary pathway. That is, an enzyme class in which it is possible to identify similarities within the catalytic core of all members of the class. Enzymes with similar activities that have evolved from convergent evolution will not necessarily share these constant residues and a model or template employing invariant amino acids as anchors would then not be possible. A variety of enzyme families are postulated to arise from divergent evolution, and thus would be expected to serve as a preferred class of enzymes for design of affector molecules within the context of the present invention. Such enzyme families include, but are not limited to, the protein kinases, phosphorylases, and several groups of proteases.

For purposes of illustration only, the present invention is described using the protein kinase family as a model system. As discussed above, these enzymes are essential for many aspects of cell regulation. Over 100 individual protein kinases have been identified. Thus, the successful design of affectors to manipulate the activity of a kinase can provides an invaluable tool for research as well as for the design of a wide variety of therapeutics and diagnostics.

Thus, for example, development of effective specific inhibitors of oncogenic kinases is believed to lead to the development of anti-neoplastic treatments. In addition, specific inhibitors of kinases involved in hormone regulation will be useful in artificially regulating the secretion and regulation of such hormones. Also, since many neuro-transmitters are regulated by kinases, development of new affectors could potentially impact on diseases of the nervous system. Further, platelet aggregation and clot formation might also be regulated through novel affectors of kinases developed through the methods of the present invention. Many other therapeutics are believed possible through the development of novel specific affector molecules.

10

5

The model system used in connection with this invention uses cAMP-dependent protein kinase together with a 20 amino acid inhibitor peptide, PKI(5-24), to establish a "lock" for specific affector design. This inhibitor is unique in that it interacts only with the cAMP-dependent protein kinase. Therefore, cocrystallization of this inhibitor with cAMP-dependent protein kinase permits the visualization of the conformation of an enzyme in association with its specific inhibitor.

15

20

The general requirements for peptide recognition by the catalytic subunit were characterized originally using synthetic peptides. From those studies evolved a general consensus sequence that includes two basic residues, typically arginine, followed by one intervening small residue preceding the site of phosphotransfer. The phosphate acceptor site, either Ser or Thr, is followed immediately by a large hydrophobic residue. The typical peptide substrate used for this kinase, based on the phosphorylation site in pyruvate kinase, is shown in 1. Replacement of the Ser with an Ala generates an inhibitor peptide with poor affinity. In contrast is the inhibitor peptide, PKI(5-24), derived from the thermostable protein kinase inhibitor (PKI). This peptide, seen in 1, encompasses the consensus sequences described above but also contains additional unique features that convey high affinity binding. The particular features that contribute to high affinity binding were defined using peptide analogues, and the most relevant ones are also indicated in 1. A general structure of the peptide in solution was deduced using circular dichroism (CD) and NMR spectroscopy. The peptide, PKI(5-24), was co-crystallized with the catalytic subunit of cAPK, and the structure of that peptide as well as its interaction with the protein are discussed here.

30

25

The folding of the polypeptide chain and the mechanism of catalysis is conserved in all protein kinases. There are 8-9 invariant residues scattered throughout the core for all protein kinases. The crystal structure reveals that most of these invariant residues are

clustered three dimensionally around the site of catalysis providing an interconnected network. The regions involved in peptide recognition extend over a wide area on the surface of the enzyme and until this structure was solved there was no understanding of the details of the peptide recognition sites. The structure of the catalytic subunit thus serves as a framework from which a template for the entire protein kinases family can be produced. This structure provides, for the first time, a true molecular basis for the design of affectors that will selectively target any given protein kinase. Thus, it is an object of this invention to provide a method for the identification and design of molecules interacting with the catalytic core of a protein kinase by preparing a template from the analysis of the catalytic core of the cAMP dependent protein kinase.

10

15

5

The ability to design affector molecules that act on a given enzyme using information obtained by X-ray crystallography is dependent on the formation of crystals of purified enzyme. Methods for crystal production vary greatly and one cannot predict how readily a given molecule or complex will crystalize. However, those skilled in the art will recognize that a variety of methods for crystallizing can be attempted for any given enzyme, and that successful crystallization can be expected of a variety of enzymes. Rational drug design additionally requires information about the interaction of a known affector in order to accurately predict a potential affector's effect on the catalytic core of the enzyme. Thus, crystals of the complex of affector molecule and enzyme together are used to gather information on the conformation of the enzyme in its inhibited conformation. Thus, in addition to information about the catalytic core of an enzyme family and the identification of additional sites adjacent to the core that permit the specific design of inhibitors, the present invention provides an improved method for the crystallization of complexes.

25

30

20

An important feature of certain aspects of this invention is the production of an enzyme/affector template. In order to generate this template, the affector chosen for production of enzyme/affector complex should have a high affinity for a particular enzyme. The initial affector molecule chosen should preferably have a Kd less than 1 µM, and more preferably less than 100 nM, in order to provide a conformation resulting from high affinity interactions. Once the specific interactions are understood it is contemplated that affector molecules having a variety of Kd ranges could be selectively designed for various purposes. Thus, in the model system chosen to illustrate this invention, PKI(5-24) is used as an affector of cAPK, with a Kd of approximately 60 nM. Those of ordinary skill in the art will recognize that other inhibitors with Kd less than 100 nM could also have been chosen to illustrate this aspect of the present invention.

For example, cGMP dependent protein kinase has an inhibitor with a Ki of approximately 6 nM, and an inhibitor for cAPK described by Ricouart et al. is characterized in the 4 nM range (J. Med. Chem., 34: 73-78, 1991). These Ki values are roughly equivalent, however, not identical, to the expected dissociation constants (Kd's).

5

The PKI(5-24) inhibitor peptide used here is highly specific and is rather large in that it extends beyond the catalytic core. Other known peptides and affector molecules for kinases are not as specific. We have discovered that the interactions beyond the catalytic core provide the high specificity of PKI(5-24) for cAPK.

10

As stated above, it is the crystallization of the enzyme with its specific inhibitor together with the analysis of the relationship of the inhibitor to both the catalytic core and to areas surrounding the core that provide data for the particular protein kinase "lock". The "lock" comprises the three dimensional structure and ionic, hydrophobic, hydrogen bonding and other interactions of the non-conserved variable residues with the specific affector structure. The lock is defined by the invariant residues of the exemplary structure. When an enzyme of the class is affected by a specific affector, the backbone atoms of these invariant residues must be in substantially the same relative coordinates in all members of the enzyme class. Thus, with knowledge of sequence information of the particular enzyme for which the affector is being designed, knowledge of the lock formed by the invariant residues can be obtained. The coordinates of the invariant residues position the variable residues of the lock in space and thereby permits the design of other specific inhibitors and affector molecules for other protein kinases.

20

15

The lock consists of the site of phosphotransfer (P site) with recognition sites for flanking sequences. The flanking sites can be identified by the number of amino acid residues separating that site from the P site. Thus, the first amino acid residue in the direction moving toward the carboxy terminus is designated P+1, and the following residues are designated P+2, P+3, P+4 and so on. Similarly, the residues on the side moving toward the amino terminus are designated P-1, P-2 and so on.

30

25

The sites for recognition of the peptide are not identical between members of the protein kinase family, and the chemical content is unique for each given protein kinase. The sequence of the given protein kinase is built into the coordinates of the C-subunit using the invariant residues. The position of these invariant residues can be identified using X-ray crystallographic data, such as the data disclosed herein in Figure 17. This data provides the coordinates for each non-hydrogen atom in cAPK. It is the locations of these invariant residues which serve to define the template common to all protein kinases. This template

can then be used to model the three dimensional coordinates of the variable as a basis to design highly specific affector molecules.

The affector molecules to be designed could be polypeptides, nucleic acids and their analogues, combinations of nucleotides and peptides, organics or any other molecule capable of specific interaction with a given enzyme. The essence of the design of a specific inhibitor for a given protein kinase is based on the three-dimensional fit of the specific inhibitor into the provided "lock", or template, provided by the known structure of cAPK.

The template defined by the invariant or other highly conserved residues can be used to define the region immediately flanking the phosphorylation site and, in addition, can incorporate more distant parts of the molecule to enhance specificity and affinity. The peptide recognition site serves in the same manner as the antigen recognition site of an antibody. This site extends over a large surface of the enzyme and provides a unique lock for the design of a wide variety of affector molecules, including both peptide and non-peptide affectors. Each particular protein kinase has a different and unique chemical content at each individual site. Thus, the "lock", is unique for each protein kinase.

The "lock" of any particular enzyme represents a topological map with defined sites, positions of which vary between members of the enzyme class. As an analogy, each kinase can be thought of as functioning in a manner similar to a specific antibody in that it recognizes only a very specific set of proteins to phosphorylate. However, each of the kinases has a conserved template, the positions of which will not substantially vary between kinases in an inhibited conformation. Thus, using computer modelling together with known sequence information regarding a particular kinase, the invariant residues of the kinase can be placed in the template conformation, and the approximate positions of the variable residues can be predicted.

The lock provides the information from which other specific affector molecules can be designed. It provides information on topology, charge interactions and the points of contact both within the catalytic core and around the core that suggest the design features important for the production or identification of novel affector molecules. Thus, the goal is to design an affector having homologous topography and charge fields that complement the catalytic core of the lock of the enzyme. Computer modelling can be used with these factors to design an affector capable of inducing a conformation where the conserved amino acid residues of the enzyme are in homologous locations to the template.

BRIEF DESCRIPTION OF AFFECTOR DESIGN

10

5

15

20

25

30

5

10

15

20

25

30

The basic steps toward achieving this invention are provided briefly here and in detail below. A class of enzymes is first identified wherein at least one enzyme of the class has a highly specific affector molecule. Then the inhibitor is tested for specificity and, preferably, the inhibitor sequence is reduced in size until a minimum sequence having the desired specificity is obtained. Sequence data from related enzymes is analyzed so that a consensus region that forms the catalytic core can be identified. Crystals of affector molecule together with the model enzyme are subjected to multiple isomorphous replacement techniques to prepare heavy atom derivatives. This permits the location of heavy atoms within the structure to be identified and additionally permits multiple diffraction patterns to be combined to deduce phase angles for calculation of the electron density of the structure. Those of ordinary skill in the art will recognize that other techniques can be used to deduce phase angles and to improve the accuracy of previously deduced phase angles.

A three-dimensional structure can be obtained from the electron density data using a computer program such as TOM/FRODO. Further, a computer program, such as X-PLOR, can be used to improve the accuracy of the initial three-dimensional structure. There are a variety of computer programs available for analyzing X-ray crystallographic data. Those used in the development of the model system for this invention are cited herein. Those of ordinary skill in the art will recognize that many other such computer programs providing similar functions could also have been used. From this data, the points of contact are identified both within the catalytic core and the surrounding region. Invariant amino acids and consensus recognition sequences are identified. The data is further analyzed against available chemical data such as NMR, CD, SANS data and other data resulting from chemical procedures. This chemical data can provide additional information for the structural model.

The coordinates of the invariant amino acids residing in the conserved catalytic core and the surrounding invariant residues in the enzyme/affector complex provide the template to be duplicated in other members of the enzyme class. The lock of the enzyme for which the affector is to be designed is then built by replacing the variable amino acids of the catalytic subunit of the template enzyme with the amino acids of the new enzyme. Any gaps in the sequence alignment between the enzyme used to generate the template and the enzyme for which the lock is being modelled generally occur within loops. These loop regions can be modelled separately using the structural data accumulated in a data bank, such as the Brookhaven data bank. The model of protein kinase can then be corrected and

refined using an energy minimization procedure and using molecular dynamics to eliminate stearic and electrostatic clashes. The resulting model of the catalytic core of the protein kinase under investigation is then inspected for amino acid content of the enzyme's surface which interacts with the proposed inhibitor.

5

In accordance with one aspect of the present invention, another member of this enzymatic class can then be analyzed in the context of this template. If the new enzyme can be crystallized, then the information obtained from the crystallization is merged with the "lock" structure. However if the new enzyme is not accessible or is not crystallizable, the enzyme can still be incorporated into the three-dimensional lock. The ability to incorporate the new enzyme into the template is dependent on the identification of conserved residues within the catalytic core of the new enzyme that are complementary to the conserved residues in the model enzyme. The template establishes the coordinates for these residues in three-dimensional space as well as providing coordinates for the three-dimensional surface of the catalytic core and adjacent regions.

15

10

For cAPK and the protein kinase family, the invariant residues are identified and summarized in the review by Hanks et al., <u>supra</u>. The template permits a comparison of the new enzyme catalytic core surface with cAPK. Residues within the catalytic core that are different from those of cAPK are studied to determine how those differences in the new enzyme might alter the surface of the core or change the structure of a new affector molecule. Recombinant cAPK can then be subjected to site-directed mutagenesis to change residues specific to cAPK into residues found in the new enzyme. This recombinant protein can be crystallized.

25

30

20

A novel affector molecule can then be synthesized that complements the electrostatic charges and topography of both the catalytic core and identified surrounding regions of interest for the new enzyme. The points of contact, hydrophobic pockets, site of phosphotransfer, topography and stearic interactions are assessed and the affector molecule can then, if necessary, be subjected to random mutagenesis or site-directed mutagenesis to improve the affector/enzyme interaction. This model affector molecule together with recombinant mutated cAPK, are tested with the target enzyme for affector activity. The affector molecule is finally tested with the native new enzyme. Fluorescent tags bound to the affector can be used to assess binding to the new enzyme in the cell. Alterations in enzyme function can be detected by gel electrophoresis and complexes of enzyme and affector can be isolated and purified for further analysis. Thus, new enzyme purification and crystallization is not required for affector design.

As discussed above, generation of new affectors is not limited to peptides. A variety of chemically synthesizable compounds can be used.

The model can be tested by a variety of methods. For example, kinetic determination of inhibition constants of novel inhibitors can be measured. Also, CD, SANS and other chemical procedures can be used to assess the extent of the conformation changes due to binding of the affector. If a mutated form of the enzyme has been prepared, cocrystallization of the affector with this mutated form can be performed and the points of contact can be determined and compared with the modelled points of contact.

ENZYME FAMILY CHOICE AND IDENTIFICATION OF CATALYTIC CORE

10

15

20

5

This invention relates particularly to enzyme families formed by divergent evolution. Once an enzyme family of interest is identified, an individual enzyme is chosen from a group of enzymes that share invariant residues within their postulated active sites.

The enzymatic or active site within a given protein kinase can be broadly identified through biochemical means. When the enzyme exists as a group of subunits, enzymatic activity is often restricted to one of those subunits. Thus, prior to performing these biochemical means, the enzymatic subunit can be purified from the holoenzyme. The active site can be further localized by systematically reducing the subunit size and assessing enzyme activity with each reduction. In one method, the various mRNA sequences encoding the related enzymes are reversely transcribed and cloned. Sequence information can then be obtained from the catalytic region for a number of enzymes of the same class. Similar amino acid residues within the catalytic subunit are aligned in order to visualize homologous regions. Invariant amino residues can be identified among the class which are either present in all known members of the class or substantially all members of the class. At least a plurality of these invariant residues are believed necessary for enzyme activity within the catalytic subunit. Thus, the invariant residues can further define the catalytic core.

25

For the protein kinase family, invariant amino acid residues are located within the catalytic core and are boxed in by a solid line in Figure 1. Figure 1 illustrates the general topological arrangement of the catalytic subunits from a number of protein kinases.

30

Figure 2 illustrates that while the catalytic regions from members of the protein kinase family share some striking similarities, the placement of this active region within the enzyme, the size of the enzyme and the regulatory regions of the enzyme vary considerably. The conserved catalytic core is denoted in Figure 2 as solid black areas and regulatory regions are cross-hatched. Additional information regarding the use of cAMP-dependent

Protein Kinase as a model for the protein kinase family can be found in a review by S. Taylor (J. Biol. Chem. 264:8443-8446, 1989.).

It is known from an analysis of the catalytic core of the protein kinase family that the core is included in a conserved 300-residue segment. Site-directed mutagenesis of recombinant enzyme sequences has been used to identify particular residues critical to enzymatic function. In the cAMP-dependent kinase an invariant lysine residue at position 72 has been shown to be important by site-directed mutagenesis and a triad of glycines is thought to be associated with ATP-binding.

The enzyme exists as a tetrameric holoenzyme composed of a dimer of regulatory subunits and two catalytic subunits. cAMP binds to the regulatory dimer yielding dissociation of the enzyme into an $R_2(cAMP)_4$ complex and two active catalytic (C) subunits. It is the active C-subunit that phosphorylates serine or threonine residues on substrates having the consensus sequence Arg-Arg-X-Ser/Thr-Leu.

AFFECTOR MOLECULE FOR TEMPLATE DEVELOPMENT

15

10

5

In a preferred form of the present invention, the enzyme used to establish the template or lock is, advantageously, a molecule that binds with high affinity to its affector, preferably with a dissociation constant less than 1 µM. For example, there are many such known affectors, such as inhibitors and activators, of various protein kinases. Kinases with a regulatory subunit are known that are inhibited by a peptide encoding the regulatory subunit binding site. Similarly, kinases that possess an autoinhibitory portion are also known. Thus, for such a kinase, this autoinhibitory region could be cleaved away from the core enzyme, purified and analyzed to provide a minimal high-affinity inhibitory sequence.

20

25

There are several inhibitors of cAMP-dependent kinase. The regulatory subunits can function as physiologic inhibitors as can the heat stable inhibitor protein (PKI). These inhibitors share a substrate-like sequence based on the arginine doublet, N-terminal to the position of the phosphorylation site in a normal substrate. Peptide fragments containing the consensus sequence bind the C-subunit in a manner analogous to a real substrate. PKI has an alanine in place of the phosphorylatable residue. While the PKI sequence is clearly inhibitory, the addition of a 15 residue stretch N-terminal to PKI increases inhibitory activity. Thus, residues external to the catalytic site are believed to be relevant in providing potent, high-affinity, inhibition and for improving the specificity of an inhibitor.

30

A protein that is a high-affinity inhibitor of an enzyme can be dissected to find a smaller fragment, if it exists, that still contains high-affinity inhibitory activity. At least three factors are useful in this dissection: 1) an ability to produce chemically defined fragments

5

10

15

20

25

30

of the larger inhibitor, either by synthesizing peptides or by cleaving the inhibitor with reagents such as cyanogen bromide or proteases, that cut at short amino acid sequences of a specific type for each reagent, 2) an ability to isolate specific fragments of the larger inhibitor from the mixture of fragments resulting from cleavage of the larger inhibitor, and 3) an ability to assay chemical species for inhibition of the enzyme of interest.

To carry out the isolation of a potential smaller inhibitory region of a larger inhibitor, one can cleave the inhibitor into fragments using a protease. Then one can separate the resulting fragments using HPLC and assay the fractions for high-affinity inhibition of the target enzyme. If no fraction is found that exhibits the desired inhibition, the cleaving reagent may have cleaved at a location that splits the inhibitory portion of the protein, destroying its ability to inhibit. In this case, it would be desirable to obtain other cleavage patterns until an inhibitory fragment is found. After obtaining the smallest possible inhibitory fragment using proteolytic cleavage of the intact inhibitor, one can chemically sequence the fragment as a step toward further defining the smallest fragment still having high-affinity inhibitory activity. With knowledge of the amino acid sequence, one can then use peptide synthesis to construct progressively shorter subsets of this fragment. These shorter subsets can then be assayed for inhibitory activity. Proceeding in this manner will thus allow definition of the smallest sequence, present in the larger inhibitor that still possesses high-affinity inhibitory activity toward the target enzyme. determining inhibition constants for tight-binding inhibitors are found in Biochem. J. 127: 321-333, 1972 by P. Henderson. Methods for determining the inhibitory region PKI(5-24), are provided by Scott et al. in Proc. Natl Acad. Sci. USA 82:4379-4383, 1985.

OBTAINING THREE DIMENSIONAL STRUCTURE DATA

In order to obtain data on the conformation of the template of the enzyme formed by binding of the affector thereto, a variety of techniques can be used. These techniques include, circular dichroism, small angle neutron scattering, diffraction methods, including any combination of multiple and single isomorphous replacement, single or multiwavelength anomalous scattering methods, molecular replacement methods maximum entropy phasing, solvent-flattening methods and so-called "direct" methods used primarily to solve small-molecule structures. However, in the preferred embodiment, X-ray crystallography is used in order to generate specific coordinates for each of the non-hydrogen atoms in the complex. Coordinates for the hydrogen atoms could additionally be obtained using neutrons. Thus, following the isolation of an exemplary protein and affector and following or during the sequence analysis of related enzymes, crystals of enzyme and affector protein are generated.

The crystals can be generated from enzyme purified from natural tissue or from enzyme generated by recombinant means. Provided below are examples pertaining to the production of crystals using the recombinant mouse C_{α} -subunit of cAMP dependent protein kinase and purified cAMP-dependent protein kinase from porcine heart. Nelson et al. describe the purification schemes for porcine heart cAMP dependent kinase (J. Biol. Chem. 256:3743, 1981.) and Slice et al. disclose the methods for the generation of recombinant mouse C_{α} -subunit in E. coli (J. Biol. Chem. 264:20940, 1989). The sequence data for cAPK was published by Uhler et al. (J. Biol. Chem. 261:15360-15363, 1986).

5

10

15

20

25

30

The steady state kinetics of the C-subunit, purified from E. coli are identical to the mammalian C-subunit, although the E. coli protein is more labile to heat denaturation. Unlike the mammalian enzyme, the recombinant C-subunit lacks a myristoyl group at its amino terminus. For a review of protein crystallography see Protein Crystallography, 1976, T. Blundell and L.N. Johnson, Academic Press, New York. Information on circular dichroism and neutron scattering is found in Biophysical Chemistry. Part II: Techniques for the Study of Biological Structure and Function, C.R. Cantor et al. (W.H. Freeman and Co., San Francisco, 1980).

Example 1

Porcine Heart C-subunit Crystal Forms

Reagents were obtained from the following sources: threo-1, 4-dimercapto-2,3-butanediol (DTT, dithiothreitol; Aldrich, Milwaukee, WI); N,N-bis(2-hydroxyethyl)glycine (Bicine; Aldrich); methanol (Fisher Scientific); ammonium acetate (Aldrich); polyethylene glycol (Dow, Midland, MI).

The peptide inhibitor PKI(5-24) was synthesized at the La Jolla Cancer Research Foundation (La Jolla, CA) and modified in our laboratory These modifications are described in detail below. The sequences of the peptide inhibitors are: (1)PKI(5-24); TTYADFIASGRTGRRNAIHD, (2)PKI(5-24), tyrosine iodinated: TTY*ADFIASGRTGRRNAIHD. The peptide sequence abbreviations follow either of the two standard abbreviation schemes for amino acids; the three letter code or the single capital letter designation. Both are standard abbreviations and are well understood by those of skill in the art.

The porcine C-subunit was purified to a single band on SDS-polyacrylamide gels and used for crystallization. Two crystal forms were prepared. Photographs of the porcine heart apoenzyme (cubic); and the porcine heart C:MgATP:PKI(5-24) ternary complex (hexagonal) are provided as Figures 18A and 18B.

5

10

15

20

25

30

The first crystal form used the hanging-drop vapor diffusion method. A drop of protein mixed with precipitating agents is suspended from a microscope cover slip and allowed to equilibrate through the gas phase against a larger reservoir.

Specifically, 30-µL drops of 3-4 mg/mL protein solution were suspended and allowed to equilibrate against approximately 1mL of reservoir solution in wells of plastic Linbro tissue culture trays over a time of several days to weeks. Both new forms, as well as the earlier P2₁ form, were grown at 4°C. The porcine heart C-subunit was concentrated to 8-12 mg/mL and subjected to a final dialysis before attempting crystallization. The specific recipe for obtaining the first new form was the following: drop- 1/3 protein in 50 mM $(NH_4)_2HPO_4$ and 5 mM 2-mercaptoethanol (pH 8.0-8.2); 1/3 150 mM NH_4CH_3COO , 50 mM (NH₄)₂HPO₄, and 10 mM dithiothreitol (DTT) (pH 8.1-8.2); and 1/3 reservoir composed of 8-9% PEG-400, 17-20% MeOH, and 10 mM DTT. Crystals of the second form were obtained from the same conditions as the first new form when the drop contained, in addition to the protein, MgATP and a 20-residue peptide inhibitor [PKI(5-24)] in the molar ratio 20:5:1:1 ATP:Mg²⁺:PKI(5-24):C-subunit. The same crystal form was subsequently grown from a drop containing 1/3 protein in 50 mM bicine, 100 mM NH₄CH₃COO, and 5 mM 2-mercaptoethanol (pH 8.3); 1/3 MgATP and PKI(5-24) in 10 mM DTT in the same ratio to protein as before; and 1/3 8 mM DTT and 8% PEG-400. The reservoir contained 8% PEG-400, 15-20% MeOH and 7mM DTT. The first new crystal form could also be grown in the presence of the Mg²⁺ and the non-hydrolyzable ATP analogue adenosine 5'\beta, γ-methylenetriphosphate (AMP-PCP). The second new crystal form, representing the ternary complex, could be grown with CoCl2 or CdCl2 substituted for MgCl2 in the crystallization. The transition from one crystal form to another caused only by addition of MgATP and the peptide inhibitor PKI(5-24) suggests that a significant conformation change may occur upon their binding.

The space groups of the new crystal forms were determined to be $P4_132$ (cubic) (Figure 18A), and $P6_122$ (hexagonal) (Figure 18B), respectively. Space groups were determined and all diffraction data were measured at the University of California, San Diego Research Resource Laboratory at 4°C using graphite-monochromated CuK_{α} X-rays from either the Mark II Elliot GX-6 rotating anode diffractometer operating at 2 kilowatts or the Mark III Rigaku RU-200 rotating anode diffractometer (available from Rigaku USA, Danvers, MA) operating at 5 kilowatts, each equipped with two Xuong-Hamlin multiwire area detectors (available from San Diego Multiwire Systems, San Diego, CA).

5

10

15

20

25

30

Preferably, area detector data collection is used. One facility offering equipment to support this data collection technique is The Resource Research Laboratory. This facility is a geographically designated, NIH supported facility to promote the use of X-ray crystallographic techniques. In speed, signal-to-noise ratio, and data precision, area detector data collection far surpasses standard diffractometer or film data collection. On the average, data collection is 50 times faster; consequently complete high resolution data sets can frequently be collected from a single crystal in one or two days. The space groups and lattice constraints of the crystal forms were determined to be the following: $P4_132$, a=b=c=169.24 Å; and $P6_122$, a=b=80.3 Å, c=293.0 Å. Calculations using an average reciprocal density of 2.7 Å³/D yield to the nearest unit 2 and 1 C-subunit monomers/asymmetric unit. The $P4_132$ form diffracts typically to 3.2 Å. Pictures of the cubic and hexagonal crystal forms can be seen in Figure 18

Because the P6₁22 crystal form had diffraction better in extent and decay characteristics than the P4₁32 form and because of the greater biochemical interest of a ternary complex, work concentrated on solving the hexagonal ternary complex crystal form. The lack of phase angles for a similar protein structure prohibited an initial structure solution for the C-subunit in the P6₁22 form using molecular replacement techniques, so a structure solution using standard multiple isomorphous replacement (MIR) techniques was attempted. Both of these are techniques known to those of ordinary skill in this art. Briefly MIR involves introduction into the space group asymmetric unit of a relatively heavy reference atom that, after being located through difference Patterson analysis, enables the needed phase angles to be determined. The reference atom can be found with 6-Å data, and with its location and generation of phase angles the fundamental crystallographic problem of a protein structure solution is solved and an electron density map can be calculated. Subsequent work on a protein structure focuses on incrementally improving the degree of detail visible in the electron density map through acquisition of higher resolution data and accompanying phase angles.

The procedure used to search for heavy-atom derivatives was to soak or co-crystallize C-subunit with heavy atoms based on the empirical success record or various heavy atom reagents and on known C-subunit chemical information, such as the availability of two free thiol groups and the obligatory use of a divalent cation in catalysis. Soaks in Au, Hg, and Pt compounds yielded precession picture diffraction changes but uninterpretable 6A difference Patterson maps. A 4.8-Å data set from a Na₂U₂O₇ soak yielded an apparent Patterson solution through examination of isomorphous difference and (1/variance)-

weighted anomalous difference Patterson maps, but the site quality was not high and attempts to reproduce or improve the soak failed. Isomorphous crystals grown with Co²⁺ or Cd²⁺ substituted for Mg²⁺ proved useless since neither metal could be located (location of Co²⁺ through Patterson analysis was improbable anyway due to its lightness), although with phases their positions could reveal metal site number and location. Co²⁺ and Cd²⁺ were chosen for co-crystallization based on their reported ability to support nucleotide binding to the C-subunit and support catalysis, although at a reduced rate.

The single most important modification in the crystallization protocol that led to the formation of crystals in a different space group was the careful selection of polyethylene glycol in combination with various low molecular-weight alcohols. Commercially available polyethylene glycol contains various contaminants that may cause problems in the achievement of stable and reproducible crystallization conditions. All commercially available polyethylene glycols (PEG) were examined with the aim of detecting the presence of ionic species.

15

10

5

The lowest level of ionic contaminants was detected in PEG manufactured by Dow Chemical. It is this PEG that was selected for further crystallization experiments. PEG from other sources appeared to be generally more contaminated and also exhibited large differences in contamination between batches. In our experiments, several molecular weights of PEG were used along with several low-molecular-weight alcohols.

20

The catalytic subunit crystallized in the hexagonal space group with the introduction of PKI(5-24) and MgATP, whereas in its apo form it crystallized in the cubic space group using otherwise identical crystallization conditions, indicates that the hexagonal crystal may arise as a result of a different conformational state of the enzyme.

Example 2

25

Mouse recombinant C-subunit Crystal Forms

One of the most promising directions for combining crystallographic methods with those of molecular biology is the development of highly effective vectors for expressing large amounts of protein for crystallization. Expression of protein in E. coli also provides a mechanism for eliminating posttranslational modifications which may hinder crystallization and in addition permits structure-function studies on mutant forms of the protein following the generation of mutant containing crystals.

30

The recombinant murine catalytic subunit, whose expression and purification was described by Slice et al., is devoid of myristic acid at the N terminus and differs by nine amino acids from the porcine heart catalytic subunit used in the earlier crystallizations. It

5

10

15

20

25

30

has been shown, however, that N-terminal myristoylation is not necessary for C-subunit function. Additional differences between the porcine heart and recombinant mouse C_{α} proteins include the presence of additional phosphorylation sites (Ser 10 and Ser 139) on the recombinant protein.

Crystals were prepared from a binary complex of the recombinant mouse C_{α} -subunit with a bound, high-affinity ($K_i = 3$ nM) inhibitor peptide. The peptide (PKI95-24) derived from the N-terminal region of the naturally occurring thermostable protein kinase inhibitor protein (PKI), is the same peptide inhibitor used for the porcine heart ternary complex crystal. The steady state kinetics of the C-subunit purified from E.coli, are indistinguishable from those of the mammalian C-subunit, although the E. coli protein is more labile to heat denaturation.

The recombinant protein was crystallized using a small variation of the porcine heart ternary complex (hexagonal) conditions. A photograph of an exemplary crystal is provided in Figure 18C. First a ternary complex was prepared with MgATP and PKI(5-24). A C:PKI(5-24) binary complex was obtained after small-angle neutron scattering experiments showed that for the recombinant mouse C-subunit, the PKI(5-24) peptide alone, without MgATP, was able to cause a significant decrease in the radius of gyration. The ternary complex crystal form diffracted to at least 2.7 Å on the Mark III and was of orthorhombic space P2₁2₁2₁. A data collection strategy following the procedure of Xuong, et al., (Acta Cryst. B41: 267, 1985 was developed. Equipment for use with this procedure is available from San Diego Multiwire Systems of San Diego, California. The procedure allowed an asymmetric unit of data to be collected in 3 ω-sweeps totaling about 140° (with appropriate choice of φ and χsettings and a crystal mounted with one of the axes parallel to the capillary axis). Data collection took about 16 h for a >90% complete 2.7-Å data set from one crystal with R_{sym} on the intensity of 4-6%; in the same period the average reflection intensity decayed approximately 15%.

Crystals were generally soaked or mounted in a stabilizing solution prepared as the crystallization drop, but with the addition of the initial reservoir MeOH percentage and the omission of C-subunit and PKI(5-24). It was discovered that Cd²⁺ could be substituted for Mg⁺² in crystal growth, as with the porcine heart ternary complex crystal. It was also discovered that elevating the MgCl₂ to ten times the starting mother liquor concentration, after crystal growth had stopped, altered the cell dimensions slightly (<1%) and resulted in a different pattern of heavy-atom binding.

Example 3

5

10

15

20

25

30

Recombinant Binary-Complex Structure Solution

The binary complex crystal was nearly isomorphous with the ternary complex crystal, differing by less than 1% along any axis and had the same space group with a = 73.62 Å, b=76.53 Å, c=80.14 Å. The asymmetric unit contains one C:PKI(5-24) complex and has a calculated solvent content of 0.53. Mercury reagents were co-crystallized with the C:PKI(5-24) complex by exposing it to 1-mM reagent for six hours, followed by dialysis to remove excess Hg reagent. Native and co-crystallized 4-(hydroxymercuri)benzoic acid (PHMB) derivative diffraction data were measured on the Mark III diffractometer. Data from two native crystals were merged to provide a more complete data set for the initial phase computation, phase extension work, and initial refinement cycles, but later refinement used data from only a single crystal that yielded better quality data. Data from two PHMB co-crystals, one with measured Bijvoet mates, were kept separate and used to compute initial phases, which were improved at constant 3.5-Å resolution using the solvent flattening approach of Wang (Methods Enzymol. 115:90, 1985), with a slightly conservative solvent content of 0.50. Additional phases from inversion of the solvent-flattened map were added incrementally using 6 additional envelopes to a final resolution of 2.7 Å. Starting with minimap α -carbon coordinates, the program TOM/FRODO (available from Christian Cambillau, University of Marseille, Marseilles, France) was used with the resulting map to model in the sequence of PKI(5-24) and 257 of the 350 residues in the C-subunit. This partial model was refined using X-PLOR (available from Axel T. Brunger, Yale University, New Haven CT), and calculated structure factors were combined with Wang structure factors in a resolution-dependent manner using equally weighted ABCD coefficients to yield improved maps. ABCD coefficients are described by Hendrickson et al. (Acta Cryst. B26: 136, 1970). Structure solution statistics are summarized in 2, and a sample of electron density of the structure determination is shown in Figure 3, described in detail below.

A number of crystal forms of the catalytic subunit of cAPK have been obtained thus far. All of the crystal forms of the different complexes of the catalytic subunit, with the exception of the monoclinic crystal of the apoenzyme, were obtained under identical crystallization conditions and these are described above. The crystals in different space groups therefore very likely result from conformational states of the enzyme. Crystals of both the binary and ternary complexes with PKI(5-24) exhibited better diffraction characteristics than crystals of the apoenzyme.

Our results also indicate that the ternary complex of the murine catalytic subunit expressed in E. coli produced a crystal of better quality than did the ternary complex of the

5

10

15

20

25

30

catalytic subunit purified from porcine heart. It is difficult to conclude whether this was due to the absence of myristic acid, the amino acid differences between the two forms, microheterogeneity in the mammalian enzyme, or a combination of these factors. It may suggest, however, that another way to improve the quality of crystals is to mutate the protein and to cocrystallize mutants if crystallization of the wild type fails.

Three factors are important for reproducible crystallization. First, the salt of the eluting buffer of the last column must be chosen carefully. Second, the purity of the protein must be verified with isoelectric focusing gels. The protein must not contain typical additives, such as glycerol and should not be frozen prior to crystallization. Third, all reagents used for crystallization must be of the highest degree of purity. If all of these conditions are met, it is possible to obtain, in identical crystallizations, three different crystal forms representing two different conformational states of the enzyme. Some of those crystals, such as those of the ternary complex with PKI(5-24), are of much better quality than the other crystals.

The current model consists of C-subunit residues 15-350 and 1-20 of PKI(5-24) and has been partially refined using X-PLOR to an R-factor of 0.195 with r.m.s.(root mean square) bond length deviation from ideality of 0.024 Å. The location of the MgATP-binding site was determined by difference Fourier synthesis with the nearly isomorphous ternary complex crystal, which showed clear density for the adenine, ribose, and α -PO₄ for the low-[Mg²⁺] ternary complex crystal. The high-[Mg²⁺] difference density showed additional features that could contain the β - and γ -PO₄ as well as metal ion(s), but an unambiguous assignment of atoms to this density could not be made.

Diffraction data is summarized in Table 1. Definitions for Table 1 are as follows: f_h , calculated heavy-atom structure factor amplitude; F_p , measured native structure factor amplitude; F_{ph} , measured derivative structure factor amplitude; ΔF_{anom} , calculated Bijvoet difference; E_{iso} , r.m.s. isomorphous lack-of-closure, E_{anom} , r.m.s. anomalous lack-of-closure; $R_c = \sum \|F_{ph} \pm F_p\| - f_h \|\sum |F_{ph} - F_p\|$.

All diffraction data were measured at 4°C using graphite-monochromated CuK_a X-rays from the Mark III Rigaku RU-200 rotating anode diffractometer equipped with two Xuong-Hamlin multiwire area detectors. Paired runs starting from settings (ω, ϕ, χ) and $(\omega, \phi + 180, -\chi)$ were used to collect Bijvoet mates (inverse beam method). Data reduction and derivative-to-native scaling were done using the UCSD area detector data processing programs (available from San Diego Multiwire Systems). $R_{\text{sym}} = \sum |I_{\text{obs}} - I_{\text{avg}}|/\sum I_{\text{avg}}$ and is shown for merged Friedel pairs.

5

10

15

20

25

30

Native-1 was used for native. Hg positions of the PHMB (hydroxymercuri)benzoic acid) co-crystal derivative were found from a difference Patterson synthesis. The heavy-atom sites in relation to the model suggest heavy-atom binding at Cys 343 (major site) and Met 58 (minor site). Positional and relative occupancy refinement of two common sites (relative occupancies 2.66, 1.87 for PHMB-1), and calculation of native phases and corresponding ABCD coefficients, were done using the program HEAVY (available from the Protein Data Bank, Brookhaven National Laboratory, Upton, NY). Solvent flattening used the Wang program package (Bi-Cheng Wang, University of Pittsburgh, Pittsburgh, PA) on imported initial ABCD coefficients and phases to 3.5 Å. Molecular envelopes were calculated with solvent content 0.50 rather than 0.53 calculated for the cell. After 3 envelopes at 3.5 Å, the resolution was extended incrementally in 6 shells to a final resolution of 2.7 Å. After convergence at 3.5 Å, the mean phase change/reflection was 36.6° and the mean figure of merit was 0.84; the map inversion R-factor was 0.181. Phase extension added 6786 phases from 5914 in the 3.5-Å starting set; 261 unobserved reflections were estimated by map inversion in the 2.7-Å set.

X-PLOR Version 2.1 was used exclusively following recommended protocols provided in the accompanying manual. Simulated annealing was performed according to a slow-cooling protocol (Brunger et al. Science 235:458-460, 1987) between either 3000K or 4000K and 300K, followed by 120 cycles of conjugate-gradient minimization. Refinement began with the partial model of Stage A to improve the coordinates for phase combination. Combined maps were calculated using the Hendrickson-Lattman scheme. Wang phases were used to 6 Å combined ones between 6 Å and 3.5 Å or 3.0 Å, and calculated phases between 3.5 or 3.0 Å and 2.7 Å. The corresponding weighted amplitudes were m_{Wang} Fo, $m_{\text{comb}}(2\text{Fo-Fc})$, and $m_{\text{sim}}(2\text{Fo-Fc})$. The model was completed by iterative refinement and building in areas not included in refinement of partial model. Refinement and R-factor (= $\sum |F_0 - F_c| / \sum F_0$) calculations used $F/\sigma > 2$ reflections (12024 Native-1 reflections; 10194 Native-2 reflections beginning with Stage B). The current R-factor of 0.195 is for 2939

and angle deviation from ideality are 0.024 Å and 4.3°.

In summary, the crystals were grown as described above using a 5-10% molar excess of PKI(5-24) and were determined to be of space group P2₁2₁2₁ with a+73.62 Å, b=76.52 Å, c=80.14 Å. The asymmetric unit contained one C:PKI(5-24) complex and had a calculated solvent content of 0.53. Native and co-crystallized 4-hydroxymercuri)benzoic acid (PHMB) derivative diffraction data were measured using Xuong-Hamlin area detectors.

atoms (no solvent atoms) with individual B-factors (r.m.s. $B = 17.6 \text{ Å}^2$). R.M.S. bond length

Data from two native crystals were merged to provide a more complete data set for the initial phase computation, phase extension work, and initial refinement cycles, but later refinement used data from only a single crystal of better quality. Data from two PHMB cocrystals, one with measured Bijvoet mates, were kept separate and used to computer initial phases, which were improved at constant 3.5-Å resolution using the solvent flattening approach of Wang with a slightly conservative solvent content of 0.50. Additional phases from inversion of the solvent-flattened map were added incrementally using 6 additional envelopes to a final resolution of 2.7 Å. Starting with a minima alpha-carbon coordinates, the program TOM/FRODO was used with the resulting map to model in the sequence of PKI(5-24) and 257 of the 350 residues in the C-subunit. This partial model was refined using X-PLOR, and calculated structure factors were combined with Wang structure factors in a resolution-dependent manner, using equally weighted ABCD coefficients to yield improved maps. Details on the combination can be found in Allured et al. Proc. Natl. Acad. Sci. USA 83:1320, 1986 and Remington et al. J. Mol. Biol. 158:111, 1982. An example of the electron density of the structure determination is shown in Figure 3.

Figure 3 is a stereo view of the electron density for the structure determination. Portions of the latest refined model of 3 β -strands are shown (top to bottom from left): 112-106, 114-121, 75-69. Figure 3A provides the 1.5- σ experimental density calculated to 2.7Å using phases after Wang improvement and extensions. Figure 3B provides the 1.5- σ (2F $_{o}$ -F $_{c}$) density calculated with 10 to 2.7-Å refined model phases. The current model consists of C-subunit residues 15-350 and 1-20 of PKI(5-24) and has been partially refined using X-PLOR to an R-factor of 0.-195 with r.m.s. bond length deviation from ideality of 0.024Å. The structure of the catalytic subunit and affector molecule are described below.

Example 4

25

30

5

10

15

20

Structural Analysis of the Catalytic Subunit

A stereo view of the backbone structure of the C-subunit with the bound peptide is shown in Figure 4. Residues 15-350 of the C-subunit and the twenty residues of PKI(5-24), in bold print, of the partially refined model are shown. The overall dimensions of the monomer $(65\text{\AA} \times 45\text{\AA} \times 45\text{\AA})$ indicate a slightly elongated molecule. Earlier hydrodynamic measurements showing a Stokes radius of 26.1Å, a frictional coefficient ration (f/f_0) of 1.19, and a radius of gyration of 20Å are consistent with this structure. The most striking feature of the overall molecular architecture is its bilobal shape with a deep cleft between the two lobes. The core of the small lobe is associated primarily with the amino-terminus, while the core of the large lobe corresponds to the C-terminal region of the protein. The cleft

between the lobes is filled by a portion of the bound inhibitor peptide in the binary complex. A difference Fourier map of the ternary complex containing both peptide and MgATP places MgATP at the base of that cleft (Figure 5). The 3.5- σ positive density contours for the (F_{ternary}-F_{binary}) difference Fourier were calculated using refined model phases in 10 to 2.7-Å range and are shown superimposed on the partially refined backbone model. Figure 5A illustrates the general localization of MgATP while Figure 5B is a close-up of difference density enclosing an oriented model of AMP.

The cleft is clearly the site of catalysis, and the peptide-induced conformational changes, observed by both SANS and circular dichroism, may be associated with a closing of this cleft. SANS established that in the absence of inhibitor and MgATP the enzyme adopted a more expanded conformation than that adopted by the enzyme in the binary complex of the enzyme and the peptide inhibitor, or the ternary complex of the enzyme, inhibitor and MgATP. This technique was also used to show that binding of the inhibitor to the enzyme did not require MgATP. Neutron scattering, in particular, established that the apo form of the enzyme adopts a more expanded conformation than the ternary complex containing MgATP and PKI(5-24). Furthermore, PKI(5-24) alone, but not MgATP, was sufficient to induce this conformational change. SANS and CD are techniques known to those of ordinary skill in this art. Accordingly, no further descriptions of these techniques are necessary.

20

15

5

10

Most of the predictions of secondary structure made prior to this crystallographic study of the C-subunit are quite inaccurate and do not correlate well with the actual structure that is provided herein. The prediction of the secondary structure by Benner et al., Adv. Enzyme Regulat. 31:121, 1991, is somewhat more accurate. It is based on chemical information and homologies within the protein kinase family and is accurate within the small lobe. However, detailed and accurate information on the structures of the protein kinase family has not been available until the discoveries presented herein.

. 30

25

The amino-terminus of the C-subunit begins with an amphipathic α -helix that lies primarily along the surface of the larger lobe. This N-terminal region differs in the recombinant and mammalian enzymes, since the recombinant protein lacks a myristoyl group at the N-terminal glycine. In the crystal structure, the first 14 amino acids are not visible. However, the surface of the enzyme in this N-terminal region is hydrophobic, suggesting a possible site for the N-terminal myristoyl moiety of the mammalian enzyme. The myristoyl group stabilizes the C-subunit but does not promote association with membranes.

5

10

15

20

25

30

The smaller lobe, consisting of residues 40 through 125, is associated primarily with the binding of the nucleotide and is characterized by a dominance of β structure. Five antiparallel β-strands comprise the core of this domain. The only helical element in the small lobe is inserted between β-strands 3 and 4 and lies on one side of the plane of the β-sheet. It consists of two parts: a two turn helix B, followed by a sharp break and a five turn helix, helix C. Based on a difference Fourier map (Fig. 5) with a ternary complex of the recombinant C-subunit containing MgATP and PKI(5-24), and supported by chemical evidence discussed below, it is clear that this small lobe is the primary site for interaction with MgATP. As seen in Figure 5, the density based on the difference map is consistent with the adenine moiety of the nucleotide oriented towards the base of the cleft beneath the β-sheet, with the phosphates facing outwards, towards the edge of the cleft. This structure is distinct from the Rossmann fold that is characteristic of many nucleotide binding proteins.

The larger lobe, in contrast, is remarkable for its predominance of helical structure. Seven helices are found in this C-terminal domain. A particularly unusual feature are the antiparallel hydrophobic helices, helix E (residues 140 through 159) and especially helix F (residues 218 through 233), that extend right through the core of this domain. The only region of β -structure in this lobe is located on the surface of the cleft at the interface between the two lobes where four antiparallel β -strands form a sheet. Most of the regions important for peptide recognition, as well as some conserved residues likely to be involved in catalysis, are located within this larger lobe.

The C-terminal 70 amino acids, residues 281 through 350, extend over a large portion of the surface of the enzyme from the bottom of the large lobe to the top of the small lobe. The part of this extended chain that passes through the region linking the two lobes appears to participate in recognition of both the peptide and the nucleotide, even though these amino acids are outside the conserved catalytic core. The other extended chain connecting the two lobes of the enzyme, residues 120 through 127, likewise, passes through this linker region between the small and large lobe and also participates in peptide recognition. Hence, this linking region consisting of both chains may contribute in part to the observed peptide-induced conformational changes described earlier. An overall two dimensional topology diagram for the C-subunit of cAPK is presented in Figure 6. Residues corresponding to the secondary structure elements are as follows: β-strands - 1:43-48, 2:57-63, 3:67-75, 4:106-111, 5:115-120, 6:1610164, 7:171-175, 8:178-183, 9:188-191; α-helices - A:15-31, B:76-82, C-84-97, D:-128-135, E:140-159, F:218-233, G-244-252, H:263-272, I:288-293, J:301-307.

10

15

20

25

30

CORRELATION OF STRUCTURAL DATA WITH CHEMICAL DATA

As discussed above, chemical data can be used to confirm the correct interpretation of the electron density map. Chemical analysis has been used as a way to obtain structural data in the absence of X-ray crystallography. Since the protein kinase family is an enzymatic group of major import, a significant body of chemical data is available. While this data cannot be used to predict a three-dimensional structure for affector modelling, it does provide a body of data that can be used to confirm and ensure the consistency of the three-dimensional structure. Thus, once a crystal structure is obtained for a model enzyme and its affector, the chemical data present in the literature can be used to examine the consistency of the model before proceeding to the design step. The three dimensional structure of the enzyme-affector complex should provide a solid explanation for the earlier chemical data. Information provided from chemical data together with structural data is used to obtain the both the template and the "lock" derived therefrom.

For example, evidence for localizing the nucleotide binding site near the aminoterminus first came from affinity labeling with an analogue of MgATP, fluorosulfonyl benzoyl adenosine (FSBA). Labeling with a hydrophobic carbodiimide, DCCD, identified two carboxyl groups near the MgATP binding site, Asp184 and Glu91, and, furthermore, established that Asp184 could be readily cross-linked to Lys72 in the apoenzyme. The structure of the binary complex without bound MgATP (Fig. 7) confirms that all three residues are localized in close proximity to one another, while the difference Fourier map with the ternary complex places these residues close to the γ-phosphate region of MgATP (see Fig. 5). Figure 7 provides stereo views of selected conserved areas. 1.5- σ (2F_o-F_c) electron density (10 to 2.7Å) is shown superimposed on the latest refined coordinates. In Figure 7A the sidechains of the invariant Lys72, Glu91, and Asp184 are shown in proximity to each other. Figure 7B shows the catalytic loop, Arg-Asp-Leu-Lys-Pro-Glu-Asn (165-171), together with part of PKI(5-24). Arg 20 of PKI(5-24) is labeled as 365. Lys72 is on βstrand 3, and Glu91 lies along the edge of the C-helix that faces the cleft. Asp184 is located on the loop connecting \beta-strands 8 and 9, and this loop also lines the cleft. All three residues are invariant in every protein kinase. Therefore these residues can be used as anchors for modeling the three dimensional structure of other protein kinases.

The MgATP binding site was defined more globally by differential labeling with acetic anhydride. By describing the reactivity of each lysine side-chain in the presence and absence of substrates, it was shown by Buechler et al., Biochemistry 28:3018-3024 (1989), that the specific protection afforded by MgATP was localized exclusively to residues in the

small lobe. In addition, to Lys72, MgATP protected Lys76 and Lys47 against modification by acetic anhydride. These protected lysines also flank the conserved glycine-rich loop that lies between β -strands 1 and 2. Based on the difference Fourier shown in Figure 5, this loop is close to the phosphates of MgATP.

Chemical studies using an affinity analogue have shown that Cys 199 is important for peptide binding. Modification of Cys 199 leads to loss of activity, and MgATP protects against inactivation. In contrast, Cys 343 can be covalently modified with no concomitant loss of activity. The structural analysis reported here indicates that Cys199 is on the surface of the cleft that interacts with the C-terminus of the inhibitor peptide, and Cys343 is on the surface of the small lobe. This distance measured between the two α -carbons of Cys199 and Cys343 in the crystal structure is 24Å. Thus, some of the chemical data is confirmed by the crystal structure.

CONSERVED REGIONS AND THEIR FUNCTIONS

5

10

15

20

25

30

The fact that all known protein kinases share a conserved catalytic core that is homologous to the C-subunit provides information that independently highlights important regions. This conserved catalytic core begins with the β-1 strand in the small lobe and extends through Arg280 in the large lobe (Hanks et al., supra) The two lobes comprising this conserved catalytic core can be seen clearly in Figure 8. Figure 8A is a space-filling model of the catalytic core (residues 40-280) shared by all protein kinases. The small lobe corresponding to the nucleotide binding fold 1 (residues (40-126); the larger lobe 2 (residues 127-280). In this model the bound peptide is not shown. Figure 8B is a diagram of the conserved catalytic core using the RIBBON program of the PAP package (J. P. Priestle, J. Appl. Cryst. 21:572, 1988 and available from the Molecular Simulation Laboratory at the University of Minnesota, Minneapolis, MN). Regions of the linear sequence noted by Hanks et al., supra, are indicated. The protein kinase having the largest insert at each position is designated using the following notation to define each insert: Gene/Protein Name: NH2-terminal C-subunit residue no. (insert length) COOH-terminal C-subunit residue no. The inserts are CDC7:64(14)65, KIN1:83(26)84, PKC-y:98(6)99, cmos:113(5)114, PDGFR:137(99)138, CDC7:196(82)197, ran+1:210(23)211, HSVK:240(11)241, CDC7:260(93)261, 7less:178(7)179. Figure 8C is identical to Figure 8A, but includes PKI(5-24) 3. Within this conserved core are nine invariant amino acids, as well as several highly conserved residues. Most of these conserved residues contribute directly to either MgATP binding or catalysis. Others, such as Arg280 and Asp208, exist as ion-pairs

and link two segments of the polypeptide chain that are widely separated in the linear sequence.

In addition, to providing information on conserved residues, sequence comparisons among protein kinases also identify inserts, sometimes quite sizable, that lie within the catalytic core. These inserts were noted by Hanks et al., <u>supra</u>, but their conformation in the overall structure of the catalytic subunit and their relationship to other regions of the catalytic core is described for the first time here. The locations of these inserts are indicated in Figure 8. All inserts invariably are located at loops on the surface of the protein and can be accommodated within the tertiary structure.

10

15

5

The structures of several important regions of the catalytic subunit are described below. Two highly conserved loops, as well as a triad of invariant charged residues, appear to be particularly important for nucleotide binding and catalysis. However the regions important for recognition of the peptide substrate are quite variable and were not available until the crystallized structure was analyzed. Predictions based on these variabilities are heretofore undescribed.

GLYCINE-RICH LOOP

The glycine-rich segment, Gly⁵⁰-Thr-Gly⁵²-Ser-Phe-Gly⁵⁵, was identified originally as part of the MgATP binding site based on its proximity to Lys72 and on differential labeling with acetic anhydride, since all of the lysines flanking this region, Lys47, Lys72, and Lys76, are protected in the presence of MgATP. The specific structural explanation obtained from crystallographic data for the protection of Lys47 is due to ionic pairing with the side chain of Glu333 while Lys76 ion-pairs with Glu346. Thus, conformation changes that occur around the glycine-rich loop as a consequence of MgATP and peptide binding are understood from the structural data in combination with known chemical data.

25

30

20

A glycine-rich motif is associated with many nucleotide binding sites, and this region has been the subject of much speculation and model building. The Rossmann fold, found in many nucleotide binding sites, contains a sheet of mostly parallel β -strands containing a glycine-rich loop. A similar motif containing a glycine-rich loop is found in other proteins such as adenylate kinase and p21 ras. The protein kinase fold found in the C-subunit and conserved in over one hundred protein kinases, does not conform to either of these motifs; it forms a unique nucleotide binding site. The uniqueness of this site is summarized as follows: (1) The glycine-rich segment lies at a sharp turn that joins two antiparallel strands at the beginning of the β -sheet. (2) The phosphate binding site is not dominated by a helix whose dipole points towards the phosphate. (3) The nucleotide does not lie along the edge

of the β -sheet. (4) An invariant Lys does not immediately follow this loop. Instead, the invariant Lys in the protein kinases, Lys72, is located in the β -3 strand and is a part of the stable scaffold of the structure. The single conserved element in each of these motifs is the glycine-rich loop whose apparent function is to serve as a phosphate anchor so that the γ - PO₄ is poised for transfer.

CATALYTIC LOOP

5

10

15

20

25

30

Another highly conserved loop in the C-subunit extends from Arg165 through Asn171 and can be termed the catalytic loop (Fig. 7B). This catalytic loop, Arg-Asp-Leu-Lys-Pro-Glu-Asn, contains 2 invariant residues, Asp166 and Asn171, and 2 highly conserved residues, Arg165 and Leu167. While the purpose of the glycine loop is to anchor the phosphate moiety and, in particular, to help position the γ -PO₄ so that it is poised for transfer, it is the catalytic loop that appears to be the central hub that communications to many different parts of the molecule. This loop not only directs the catalytic event, but also guides the peptide into its proper orientation so that catalysis can occur. The loop itself and, in particular, the residues that are important for catalysis are highly conserved, while the parts of the loop that direct the peptide binding are not.

Asp166 is one of 4 invariant carboxyl groups in the protein kinase family. It is the only one that is oriented towards the Ala side chain at the pseudo-phosphorylation site in the bound inhibitor peptide. Asp166 most likely functions as a catalytic base. Catalysis is thought to occur as a direct in-line transfer without an enzyme bound phospho-intermediate. INTERDOMAIN CONTACTS:

The triad composed of the side chains of Lys72, Asp184, and Glu91, shown in Figure 7A, is conserved in every protein kinase and is close to the γ -PO₄ of MgATP. Asp184 was a candidate for the catalytic base; however, the structure indicates that a more plausible role is participation in the chelation of Mg²⁺ in the MgATP complex. The side chain of Asp 184 also comes within 4-5Å of the side chain of Asn171. This cluster, Asp184, Asn171, and Asp166, thus forms a second triad of invariant amino acids. Asp184, being a component of both triads, has the potential to shuttle between the two conserved loops, the glycine-rich loop in the small lobe and the catalytic loop in the larger lobe. Hence, if the position of Asp184 changes following the binding of MgATP, as it probably will given its location in the structure relative to the MgATP binding site, the consequences will have a direct impact on both conserved loops. If, for example, Asp184 participates in the chelation of Mg⁺², its negative charge would be sequestered from the catalytic loop, thus allowing the other

10

15

20

25

30

residues to rearrange in order to maximize the nucleophilicity of the serine hydroxyl moiety that is poised to receive the phosphate from ATP.

This is the first protein kinase structure to be reported. The protein kinases represent a large family of over 100 enzymes that includes growth factor receptors as well as many oncoproteins. In spite of the tremendous diversity of these enzymes, all share a conserved catalytic core that retains the same essential features of secondary and tertiary structure and the same general mechanism of catalysis. The essential hallmarks of this conserved core include: (1) two lobes with a cleft between that is occupied by the substrates, (2) a unique nucleotide binding fold dominated by β -structure, (3) a largely helical domain associated with peptide binding and catalysis, (4) two β -sheets converging at the active site near the domain interface, and (5) two conserved loops, one in each lobe, that converge at the active site. In marked contrast to these conserved features shared by all protein kinases, recognition of the peptide by the catalytic subunit involves non-conserved amino acids, and the peptide binding sites extend over diverse and widely separated regions on the surface of the enzyme. The detailed structure of the bound inhibitor peptide and its specific interactions with the catalytic subunit are described below.

Affector binding site data may incorporate information derived from several experimental avenues. In addition, to crystallographic studies, substrate analogues provide insights into the specific features of a given substrate that are important for recognition. Chemical approaches such as affinity labeling and group specific labeling can identify regions and specific residues that are in close proximity to substrates. Crystallographic studies can include a structural analysis of the apoenzyme, i.e. the structure of the enzyme without other associated molecules. However, more importantly, crystallographic studies of co-crystals of the enzyme with bound substrates or affectors are provided, so that the precise features of the active site can be defined.

Thus, in the model system of the present invention, crystals of the cAMP-dependent protein kinase C-subunit/PKI(5-24) were obtained and structural data derived therefrom. This structure of the catalytic subunit is presented as example 4. The inhibitor peptide PKI(5-24) is a fragment of the heat stable protein kinase inhibitor. Additional information about this inhibitor can be found in a publication by H.-C. Cheng et al. (Biochem J. 231:655-661, 1986). This peptide includes the consensus features common to all peptide substrates and inhibitors of cAMP-dependent protein kinase. In addition, it contains other features that convey unique high affinity binding. The crystals of complexed enzyme and inhibitor

10

15

20

25

30

provide insight into the guidelines necessary for designing affector molecules for other protein kinases.

The general requirements for peptide recognition by the catalytic subunit were characterized originally using synthetic peptides. These studies found a consensus sequence that includes two basic residues, typically arginines, followed by one intervening small residue preceding the site of phosphotransfer. The phosphate acceptor site, either Ser or Thr, is followed immediately by a large hydrophobic residue. The typical peptide substrate used for this kinase, based on the phosphorylation site in pyruvate kinase, is shown in Table 2. Replacement of the Ser with an Ala generates an inhibitor peptide with poor affinity. In contrast is the inhibitor peptide, PKI(5-24), derived from the thermostable protein kinase inhibitor (PKI). This peptide, seen in 2, encompasses the consensus sequences described above but also contains additional unique features that convey high affinity binding. The particular features that contribute to high affinity binding were defined using peptide analogues, and the most important ones are also indicated in 2. This peptide, PKI(5-24), was co-crystallized with the catalytic subunit, and the structure of that peptide as well as its interaction with the protein are discussed below.

A schematic of substrate and inhibitor peptides of cAMP-dependent protein kinase are provided in Table 2. The nomenclature used for the peptides designates the phosphorylation site or pseudophosphorylation site residue as P. In the case of substrates, P will be Ser or Thr; in the case of PKI(5-24), P is Ala. The residues flanking this site are designated as P+1, P-1, etc. as indicated. This nomenclature provides a common frame of reference for all peptide substrates and inhibitors and can be invoked readily for every protein kinase.

The Ser peptide is based on the *in vivo* phosphorylation site in pyruvate kinase. Residues shown to be important for peptide recognition are shaded and were identified using synthetic peptide analogues of the Ser peptide and of PKI. Procedures for determining which residues are important for peptide recognition using peptide analogues can be found in articles by Glass et al. and Kemp et al. (J. Biol. Chem. 262:8802-8810, 1989 and J. Biol. Chem. 252:4888-4894, 1977 respectively.)

Example 5

Conformation Determination of the Bound Inhibitor

The conformation of bound PKI(5-24) is shown in Figure 9. Backbone C and N atoms are shown in bold. Residues particularly important for binding are labelled according to the nomenclature of 2. The amino-terminus extending from the P-16 Thr through the P-8

Ala forms an amphipathic α -helix. This helix is followed by a turn flanked by glycines at the P-7 and P-4 positions. The glycines may be important for accommodating the turn or for providing flexibility to facilitate binding of the Arg that follows each Gly. The remainder of the peptide is in an extended conformation, and the density corresponding to the region at the C-terminus, the P+2 Asp and the P+3 His, is not well defined.

The catalytic subunit itself consists of 2 lobes - a smaller lobe, associated primarily with MgATP binding, and a larger lobe. Nearly all of the features necessary for peptide recognition are found within the larger lobe, although the specific residues involved are widely dispersed both in the linear sequence and on the surface of the enzyme. The extended portion of the peptide that includes the consensus region for recognition of all substrates and inhibitors lies along the surface of the cleft corresponding to the larger lobe. The helical segment of the peptide is amphipathic, and its hydrophobic side lies in a hydrophobic pocket on the surface of the large lobe. The specific interactions of the peptide with the protein can be described by (i) the interactions that account for the unique highly affinity binding of PKI and (ii) by the features of the protein that are important for recognizing the consensus sequence common to both the inhibitors and substrate.

HIGH AFFINITY BINDING SITE

5

10

15

20

25

30

Based on the crystal structure, the high affinity binding attributed to the N-terminus of PKI(5-24) is dominated by hydrophobic interactions involving primarily the phenylalanine side chain at the P-11 position. Glass et al. showed that a replacement of this Phe with an Ala caused a 100-fold increase in K_i while replacement with 1'-napthylalanine, a residue that is considerably larger and more hydrophobic than Phe, actually decreased the K_i by 4-fold. Figure 10 illustrates the high affinity binding site interactions. A hydrophobic pocket on the surface of the C-subunit nicely complements the hydrophobic face of the helix in the inhibitor peptide. This hydrophobic pocket is lined by residues 235 through 239, Tyr-Pro-Pro-Phe-Phe, with the phenyl ring in the inhibitor peptide sandwiched between the side chains of Tyr235 and Phe239. In Figure 10A the P-11 Phe ring (numbered 356) is shown in relation to three nearby hydrophobic residues of the C-subunit: Tyr 235, Pro 236, and Phe 239. Figure 10B illustrates the interactions of the P-11 Phe and P-6 Arg sidechains with the C-subunit. Distances between charged-residue sidechain atoms <3.5Å apart are indicated by thin connecting lines. Based on the structure, the Tyr at the P-14 position is not essential for this hydrophobic interaction. In addition, to the hydrophobic interactions associated with the helix, the orientation of the high affinity binding region PKI(5-24) is fixed by the ionic

contacts involving the P-6 Arg. Two nitrogens in the guanidine side chain of this Arg undergo ion-pairing with the two oxygens of the carboxyl group of Glu203.

CONSENSUS RECOGNITION SITE

5

10

15

20

25

30

Figure 11 summarizes the interactions that contribute to recognition of the common consensus region of the inhibitor peptide. The interactions of the P-3 and P-2 Arg residues and the P+1 Ile residue with C-subunit residues are shown. Lines are drawn between charged-residue sidechain atoms <3.5Å apart. The P+1 Ile sidechain projects into the hydrophobic area formed by Leu 198, Pro 202, and Leu 205. Electrostatic interactions dominate the portion of the peptide proximal to the site of phosphotransfer, while hydrophobic interactions dominate the C-terminal region distal to the phosphotransfer site.

Two important requirements for peptide recognition by cAPK are basic residues at the P-3 and P-2 positions. Others have shown that replacing either Arg in the Ser peptide substrate leads to a 16-400-fold increase in K_m , even when the Arg is replaced with a Lys. The environment flanking the P-3 and P-2 arginines explains these results since each Arg interacts with more than one carboxyl side chain.

Table 3 provides a listing of the amino acid residues present at the various points of contact between PKI(5-24) and two protein kinases, cAPK and casein kinase II (CKII). It can be seen from Figure 11 and Table 3 that in the C subunit of cAPK, that those residues lining the p+1 site are very hydrophobic and provide a pocket for the hydrophobic p+1 residues. In CKII, the residues lining this pocket are all basic or positively charged. This basic pocket compliments an acidic residue at the p+1 position and this is consistent with the known specificity of CKII, i.e. CKII prefers acidic groups at the p+1 position.

Figure 12 provides information on the consensus recognition site binding interactions. The electron density corresponding to the anionic P-3 site is shown in Figure 12A. Residue numbers 361, 364, 365, and 368 correspond respectively to PKI(5-24) P-6, P-3, P-2, and P+1 residues. The electron density of the P-3 Arg sidechain tip is shown in proximity to Thr 51 carbonyl in the glycine-rich loop, and Glu 127 and Glu 331 sidechain carboxylates of the domain-linking region. In Figure 12B the 1- σ (2F₀-F_c) electron density of the P-2 Arg sidechain is shown in proximity to sidechain carboxylates of Glu 170 of the catalytic loop and Glu 230; the P-6 Arg sidechain is shown near sidechain carboxylate of Glu 203. In Figure 12C the 1.5- σ (2F₀F_c) electron density of the P+1 Ile sidechain is shown projecting into a hydrophobic pocket comprised of residues Leu198, Pro202, and Leu205. The side chain of this P-3 Arg interacts with Glu127. The carboxyl side chain of Glu331 also is approximately 3A from the guanidinium nitrogens. The tip of Asp329 is approximately

 5\AA away. Thus, the position of the guanidinium moiety is fixed. In addition, the side chain of the P-3 Arg comes close to the backbone carbonyl of Thr51 in the glycine-rich loop and to the hydroxyls of the ribose ring. The side chain of Glu333 lies close to Lys47 in β -strand 1, and the side chain Glu334 is approximately 3\AA from the hydroxyl group of Thr48.

5

10

15

The P-2 site is also very anionic, and this Arg, likewise, interacts with more than one carboxyl group. As indicated in Figures 11 and 12B, the e-nitrogen forms an ion-pair with Glu170, while one of the terminal nitrogens interacts with Glu230. Glu 203 also comes close to this guanidinium side chain; however, its interaction with the P-6 Arg is dominant. In the absence of an Arg at the P-6 position, Glu230 may ion-pair with the P-2 Arg. Unlike the P-3 recognition site, all of the carboxyl groups at the P-2 site are an integral part of the large lobe.

DISTAL HYDROPHOBIC SITE (P+1)

Peptide analogue studies of others predicted a hydrophobic requirement at the P+1 position since replacement of the Leu with Gly in PKI(5-24) caused a 150-fold increase in K_m. The reasons for this requirement are now clear from the structure (Figs. 11 and 12C). Leu198, Pro202, and Leu205 form a hydrophobic groove that surrounds the Ile side chain. This hydrophobic region that constitutes the P+1 site lies at the edge of the cleft and is likely important for proper orientation of the actual site of phosphotransfer at the P position. In the binary complex this region begins to align in an antiparallel β-like configuration with the carbonyl of the P+1 Ile coming less than 4Å from the backbone amide of Gly200 and the carbonyl of Gly200 coming within approximately 3Å of the backbone amide of this Ile at the P+1 position (Fig. 13B). Substitution of a Pro for Leu at the P+1 position in the Ser peptide (2) yields an extremely poor substrate. Nevertheless, a depsipeptide analogue of this peptide lacking an amide proton at this P+1 site is still a good substrate for the catalytic subunit.

25

30

20

Figure 13 illustrates the catalytic site area. Residue numbers 364 and 367 correspond to the P-3 Arg and the P Ala. $1.5-\sigma$ ($2F_0$. F_c) electron density is shown in all cases. Figure 13A provides the site of catalysis together with the possible catalytic base sidechain of Asp 166 near the β -C of the P Ala. Thr 51 of the glycine-rich loop is shown near the P-3 Arg sidechain, and hydrophobic sidechains of residues Phe 54 (at the loop apex) and Phe 187 are shown near the site of phosphotransfer. The addition of a hydroxyl group would place the side chain of the residue at the P position close enough for a direct transfer of the γ -phosphate for MgATP. The side chain of the P-1 Asn also interacts with the glycine-rich loop as shown in Fig. 13A. Figure 13B diagrams the consensus recognition

site residues Arg-Arg-Asn-Ala-Ile together with the glycine-rich phosphate anchor loop to the left and residues 198-202 to the right. The term "residue" is here used interchangeably with amino acid. The carbonyl of Gly 200 can be seen pointing to the amide N of the P+1 Ile. Figure 13C shows the phosphate of Thr 197 buried among sidechains of cationic and H-bonding residues (His 87, Arg 165, Lys 189, Thr 195). Cys 199 is also shown nearby.

Thr197, one of the two stable phosphorylation sites in this enzyme, also flanks the P+1 site. Multiple electrostatic interactions, seen in Figure 13C, hold this PO₄ in place and account for its resistance to removal by phosphatases. Fixing this phosphate moiety contributes conformational stability, not only to Thr197 but also to the adjacent hydrophobic residues important for recognition at the P+1 site and for the proper orientation of the site of phosphotransfer. Based on the crystal structure, this anionic group appears to be important for the final correct assembly of the structure.

CORRELATION WITH EXPERIMENTAL PREDICTIONS

5

10

15

20

25

30

Several chemical approaches identified amino acid side chains that contribute to peptide recognition. Differential labeling with a water soluble carbodiimide, 1-ethyl-3(3-dimethylaminopropyl) carbodiimide (EDC), targeted solvent-accessible carboxyl groups that were accessible in the free C-subunit but protected in the presence of substrate (Buechler et al. Biochemistry 29:1937-1943, 1990). Two regions were identified using this approach. Glu170 was very reactive in the absence of peptide, but fully protected in the presence of peptide. The other region was the cluster of carboxyl groups near the C-terminus, Asp³²⁸-Asp-Tyr-Glu-Glu-Glu-Glu³³⁴. As indicated in Figure 12, Glu170 interacts with the P-2 Arg while the cluster of carboxyl groups flanks the P-3 site.

The crystal structure localized Cys199 close to the peptide recognition site and to the γ -PO₄ subsite of ATP. In the binary complex, Cys 199 does not appear to participate in peptide binding other than to contribute to the general hydrophobic environment around the P+1 site.

CONFORMATIONAL CHANGES ASSOCIATED WITH PEPTIDE BINDING

Substrate-induced conformational changes are associated with peptide binding to the catalytic subunit. Global changes in conformation, first observed using circular dichroism, showed both a loss of alpha-helical content and an increase in beta structure following peptide binding. A global change in shape also was observed using low angle neutron scattering. These results demonstrated a reduction of the radius of gyration (Rg) following substrate binding and furthermore established that the inhibitor peptide alone, but not MgATP, was sufficient to cause the reduction in Rg. The substrate-induced reduction in Rg

10

15

20

25

30

indicates that the apoenzyme corresponds to an open configuration of the protein while the binary and ternary complex represent a closed configuration.

The recognition of the peptide by the catalytic subunit is believed to be a multistep process. The initial step, associated with a loss in α -helical structure, was induced by both the substrate and inhibitor heptapeptides shown in 2. The second step, presumably corresponding to the final orientation of the peptide into the correct position at the active site, was associated with an increase in β -structure and could only be accomplished with the substrate peptide, not by the Ala peptide inhibitor. This increase in β -structure is probably due, in part, to the P+1 region of the peptide interacting the protein. Understanding these substrate-induced conformational changes will eventually require a detailed comparison of the apoenzyme structure with binary and ternary complexes containing inhibitors and substrate peptides both in the presence and absence of MgATP.

The peptide-induced conformational changes in catalytic subunit may reflect a closing of the cleft and probably involve the region linking the small and large lobes as well. This linker region consists of two chains: residues 123 through 127 and a highly acidic segment, residues 328-334. The P-3 peptide binding site is the only region of the inhibitor peptide that interacts directly with both of these extended chains that link the two lobes. One anionic group at the P-3 site is Glu127 and the other is Glu331. Since several of the carboxy groups in the C-terminal linking chain also interact with portions of the nucleotide binding site, even in the binary complex, this P-3 residue may contribute to the substrate-induced conformational changes.

CONSERVED AND VARIABLE SITES IN PROTEIN KINASES

The recognition of a protein substrate by the catalytic subunit is not unlike the recognition of a protein antigen by the variable domain of an immunoglobulin. The binding sites of both structures are dominated by interfacing β -sheets surrounded by loops that participate in recognition of the protein. The catalytic subunit also has helical regions, but it is the β -sheets that converge at the active site and it is the loops that play the dominant role in peptide recognition and catalysis. One β -sheet comes from the small lobe and the other from the large lobe. These two sheets are sandwiched together at the cleft. In the case of protein kinases, two of the loops are essential for catalysis and are highly conserved, unlike the immunoglobulins, whose function is only to bind antigens.

The two essential conserved loops that assemble at the site of catalysis in the catalytic subunit, seen in Figures 14 and 15, are the glycine-rich loop in the small lobe and the catalytic loop in the large lobe. Both lie on the surface that lines the cleft between the

5

10

15

20

25

30

two lobes. The glycine-rich loop serves as an anchor for the phosphates of MgATP, whereas the catalytic loop is essential for peptide binding and catalysis. Key features of the active site of the catalytic subunit are shown in Figure 14. Nine of the amino acids that are nearly invariant in all protein kinases are indicated. Gly186, another invariant residue, is not shown. The alpha carbons are in black, oxygens dotted, and nitrogens in horizontal hatching. The position of the phosphorylation site at Thr197 is indicated by vertical hatching. The portion of the active site associated with the small lobe is shaded and includes three of the invariant amino acids, Gly 52, Lys72, and Glu91. The remaining six are located in the large lobe. Residues close enough for hydrogen bonding or ion pairing are indicated by a dashed line while residues within 4-5 Å of one another are connected by a dotted line. As seen in Figure 14, seven of the nine invariant amino acids conserved in all protein kinases are located here, either in the loops themselves or connecting directly with loop residues. The single invariant glycine, Gly52, lies in the phosphate anchoring loop. The proposed catalytic base, Asp166, as well as Asn171, are in the catalytic loop. It is remarkable how thoroughly interconnected this region is with multiple ion pairs providing a finely tuned scaffolding for communication at the active site.

The three invariant residues in the small lobe all participate in nucleotide binding. Unlike Gly52, which is part of a flexible loop, both Lys72 and Glu91 are anchored to defined parts of the secondary structure - Lys72 to β -strand 3 and Glu91 to the C-helix. The difference Fourier map shows the phosphate density near these residues, with the presumed γ -PO₄ density close enough to the P Ala C_{β} for phosphotransfer were it a Ser[Thr] and indicates that these residues play a key role in the recognition of the phosphates of MgATP.

In the catalytic loop the two invariant residues, Asp166 and Asn171, interact with each other. Not only are their side chains close, but, more importantly, the nitrogen in the amide side chain of Asn171 is less than 3Å from the backbone carbonyl of Asp166. One additional nearly invariant residue, Asp220, contributes directly to stabilization of the catalytic loop. The two oxygens of this carboxylate come with hydrogen bonding distance of the backbone carbonyl and amide of residue 164 that immediately precedes the loop. The interaction of the catalytic loop with a conserved residue that lies deep within the large lobe fixes the loop from one side while peptide binding and interactions with the small lobe fix it from the opposite direction. As seen in Figure 13A and 13B, the consensus region of the peptide is sandwiched between the P+1 site on one side and the glycine-rich loop on the other side.

10

15

20

25

30

Of all the invariant residues, Asp184 is the only one that appears to communicate with both the small lobe and the large lobe. In the binary complex, it is most closely associated with Lys72, but it is also only 4-5Å from Asn171 and Asp166 in the catalytic loop. Although not shown in Figure 14, Asp184 is itself part of a tight turn with the carboxylate located within hydrogen bonding distance of the backbone amide of Gly186, another invariant residue. This entire segment, Asp166-Phe-Gly, is highly conserved in all protein kinases, and hydrogen bonding to stabilize the turn is probably conserved as well. Asp184 certainly has the potential to shuttle between the two conserved loops, and it is anticipated that the contacts of Asp184 will differ somewhat in both the apoenzyme and in the ternary complex containing bound MgATP as well as peptide. If Asp184 participates in the chelation of Mg2+, as disclosed above, then this charge will be sequestered from the immediate environment of the catalytic loop. Other residues close to the conserved residues in the catalytic loop in the binary complex are Tyr164 and Lys 168. The Tyr 164 side chain is less than 3Å from the side chain nitrogen of Asn171, and the Lys168 side chain comes close to the carboxylate of Asp166. Either Tyr or His, another good hydrogen-bonding residue, is always found at position 164, so this contact can also be conserved. Any significant change in the position of Asp184 will likely change the environment of the catalytic loop. Asp184, as well as Asn171 and Asp166, have also been identified as a sequence motif associated with many phosphotransferases, and this may represent a common mechanism among protein kinases.

The versatility and importance of the catalytic loop is highlighted not only by the conserved networking of essential amino acids at the active site, but also by the special ways in which this conserved network communicates with the variable residues that compose the peptide binding sites. This communication specifically involves loop residues that are not highly conserved. Glu170, for example, contributes directly to the anionic P-2 site. Thr201 in the P+1 site, on the other hand, comes very close to the side chain of Asp166. These two particular regions of contact involving the peptide binding site and the catalytic loop, Lys168 -Pro-Glu and Thr201-Pro-Glu-Tyr-Leu-Ala-Pro-Glu, contain sequences that differ characteristically between the kinases that transfer phosphate to Ser/Thr and those that transfer phosphate to tyrosine (Hanks et al., supra).

Arg165 is actually highly conserved in most protein kinases, and it connects in a unique way with the P+1 peptide binding site. Specifically, it points towards the phosphothreonine and helps to fix that phosphate so that the hydrophobic groove that follows and provides a pocket for the side chain of the P+1 residue is firmly positioned

(Figure 13C). This is an autophosphorylation site, and it is the only phosphorylation site in the catalytic subunit that could conceivably result from an intramolecular autophosphorylation. Chemical analysis has shown that this phosphate is very resistant to removal by phosphatases and based on this crystallographic data, appears to contribute to the final conformation stability of the enzyme. It should be emphasized as well that a phosphorylation site in this region of the protein is not a conserved feature of all protein kinases. Some kinases such as pp60^{c-src}, a protooncogene whose viral counterpart is found in Rous Sarcoma Virus, do have an autophosphorylation site nearby, but many others do not. Whether the catalytic loop communicates in unique ways with other autophosphorylation sites in other protein kinases remains to be established.

The two invariant residues that are most distant from the active site are Asp208 and Arg280. These residues constitute a conserved ion pair that lies just beneath the P+1 site and appears to stabilize a very hydrophobic region that buttresses the P+1 peptide binding site.

15

20

25

30

10

5

Figure 15 is a schematic of the conserved and variable regions of the catalytic subunit. The ribbon diagram depicts the folding of the catalytic subunit. Conserved regions include two loops - the glycine-rich loop and the catalytic loop - and are indicated. The variable peptide binding sites are shown as solid areas. Invariant amino acids Gly52, Lys72, Glu91, Asp166, Asn171, Asp184, Glu208, Asp220, and Arg280 are indicated by a large dot and are numbered. Dashed lines indicate residues that are close enough to pair, while the dotted line extends from Arg165 to the Thr197. Several points should be emphasized regarding the recognition of a peptide or protein substrate by the catalytic subunit. First is the number of sites and their diversity. Some of these peripheral peptide recognition sites are hydrophilic and highly charged; others are hydrophobic. As seen in Figure 15, most are found within the large lobe of the catalytic core shared by all protein kinases, but some also lie outside of this boundary. A second observation is that the requirements for recognition at the consensus site are not absolute. A comparison of in vivo phosphorylation sites reveals that the actual residues at each site vary somewhat as does the spacing between the positively charge side chains and the site of phosphotransfer. Thus, even in the consensus region, some variability can be tolerated. A third point is the potential for variability in recognition of different inhibitor, and presumably substrate, proteins that bind with a high affinity to the catalytic subunit. Most of the features essential for the high affinity recognition of PKI are apparent from this structure of the binary complex. The regulatory subunit, however, also binds to the C-subunit with a subnanomolar affinity in the absence

10

15

20

25

30

of cAMP. The consensus region, P-3 through P+1, is shared by both molecules. However, the R-subunit, cleaved at the P-5 position, still retains is high affinity binding for the C-subunit. In addition, the P-16 to P+1 region of the R^I-subunit is Pro-Pro-Pro-Pro-Asn-Pro-Val-Val-Lys-Gly-Arg-Arg-Arg-Arg-Gly-Ala-Ile, and this certainly cannot conform to the helical motif that dominates the corresponding region of PKI(5-24). Hence, an amphipathic helix is not required for the high affinity binding of the regulatory subunit. Instead, the residues that contribute to the high affinity binding of the regulatory subunit, specifically, must lie beyond the P+3 position and may complement a different portion of the surface of the C-subunit. This variability presumably can also extend to protein substrates where the catalytic subunit may recognize unique sequences that lie outside the consensus site.

Unlike the conserved residues that are invariant in all protein kinases, the sites involved in peptide recognition differ for each kinase. About 30% have some general similarities to cAPK. Others are quite different. However the template allows us to predict the specificity of each contact point. Figure 16 provides the sequence of PKI(5-24) and illustrates the distances between the points of contact and the catalytic site in three-dimensional space as measured from the template. The P site or site of catalysis is denoted by an arrow. Asterisks designate sites particularly important for the high affinity binding of PKI(5-24). Recognition sites essential for PKI binding to other substrates are denoted as labelled archways p+1, p-2, p-3, p-6, and p-11. All of the distances, with the exception of the p+1 site, are greater than 5 Å. That positions 5 Å or greater from the site of catalysis are important for inhibitor specificity have heretofore been undisclosed.

The identification of the subsites that are important to maintaining the specificity of the affector molecule interaction and provide Kd less than 100nM facilitates the design of other inhibitors. PKI(5-24) can be used as a scaffold for molding new inhibitors, and in addition once the electrochemical interactions are understood from an analysis of the three dimensional template, other affectors that are not peptides can additionally be identified. Thus, affectors could come from a group including but not necessarily limited to peptides, polypeptides, unmodified molecules existing in nature, synthetic molecules, nucleic acids, polymers, organics, or hydrocarbons. Molecules that exist in nature and that are known to interact with enzymes could be modified to produce affector molecules. Examples from this group include antibodies, antibiotics, protein, other enzymes, lipids, polysaccharides, saccharides and vitamins. Thus, inhibitors can be designed that utilize both conserved and nonconserved points of contact.

WO 93/02209

5

10

15

20

25

30

-50-

PCT/US92/06137

The invariant residues within the protein kinase family and specifically, cAPK, are used to apply the template and its coordinates to other protein kinases. There are 8-9 invariant residues for the protein kinase family. Other families may have differing numbers of invariant residues. Table 4 list the invariant residues and the distances between these resides. The distances are calculated between α -carbons. The distances between residues 52, 72 and 91 are expected to remain close to constant since these residues are all in the amino-terminal domain of the protein. Similarly, the distances between residues 166, 171, 184, 186, 208 and 280 would be expected to remain constant due to their being in the carboxy-terminal domain. Motion of the amino-terminal domain relative to the carboxy terminal domain is expected to change the distances between residues in different domains.

The distances calculated in Table 4 help form the model template since these threedimensional positions are taken from the crystal diffraction patterns and help to define a conserved shape for the protein kinase family catalytic core. A point of contact is defined herein to occur at the invariant residues and is additionally defined as a point of close spatial approximation between the atoms of the residues within or around the catalytic core and the atoms of the affector. These points of contact affect the specificity and the Kd of the enzyme/affector interaction.

The template is best described by Figure 11. The coordinates for the template listed in Table 4 and Figure 17 provide the spatial characteristics that permit one of skill in the art to input the template structure into a computer program and perform the invention disclosed herein. While the coordinates together define a three-dimensional surface that permits visualization of the catalytic site, there are invariant residues that establish important foci within the structure.

Lys 72 is invariable within the catalytic site for the protein kinase family and is an anchor for superimposing other protein kinases onto the template. Asp 166 can additionally be a second important anchor. Similarly the other invariant positions likewise have importance for fitting other kinases. A combination of the coordinates with the invariant residue positions allows important regions within and around the catalytic site to be visualized. From a study of the interaction of cAPK with PKI(5-24), important hydrophobic and ionic interactions can be analyzed. When a new enzyme is superimposed onto these coordinates these hydrophobic and ionic interactions are assessed with PKI(5-24). It is then possible to study what changes can be made to PKI(5-24) to model a new affector. A study of the residue sidechains and the charge distribution within the site is used to fine tune the new affector.

10

15

Any protein kinase having homology in and around the catalytic site with cAPK can be used to design specific affector molecules. Hanks et al. provides a list with homologous residues highlighted. Many growth factor receptors have protein kinase activities. These include but are not limited to platelet-derived growth factor, colony stimulating factor, the insulin receptor family and epidermal growth factor. Protein kinases are involved in hematopoiesis and lymphopoiesis. Some, like myosin light chain kinase, are calcium-calmodulin dependent, and further, a variety of protein kinases are oncogenic products. These include but are not limited to viral and cellular homologues of src, mos, abl, Neu, Fgr, and Yes. Any of these kinases as well as others fitting the characteristics disclosed herein could be used in this invention to produce specific affector molecules.

The phosphorylation target sequences are available for a variety of protein kinases. These include phosphorylatable amino acids with their surrounding residues. For some kinases this will provide a good starting point for inhibitor design. Other protein kinases have a regulatory subunit associated with the catalytic subunit in the inactive form. The binding sequences with the regulatory subunits are other starting points for affector molecule design. Additionally, there are a group of protein kinases that have a regulatory domain. This domain binds the catalytic site when the enzyme is inactive. Binding of an exogenous molecule changes the kinase conformation such that the regulatory domain no longer binds. A review by Pearson et al. provides a table of protein kinase phosphorylation site sequences (Methods in Enzymology Vol. 200, 1991 in press).

Once a template is created there are several options available for designing an affector molecule and these were outlined in the section above entitled "Brief Description of Affector Design."

EXAMPLE 6

25

20

Inhibitor design for pp60c-src without

pp60^{c-src} purification

pp60^{c-src} is the proto-oncogene homologue of the src protein kinase from Rous Sarcoma Virus. The protein causes unrestrained cell proliferation. In this example, the invariant residues for pp60^{c-src} are identified with a star and in bold below:

30

 ${\tt ESLRLEVKLGQGCFGEVWMGTWNGTTRVAIKTLKPGTMSPEAFLQEAQVMKKL} \\ {\tt RHEKLV}$

35 QLYAVVSEEPIYIVTEYMSKGSLLDFLKGETGKYLRLPQLVDMAAQIASGMAYBE

-52-

VHRDLRAANILVGENLVCKVADFGLARLIEDNEYTARQGAKFPIKWTAPEAALY GRFTI

 ${\tt KSDVWSFGILLTELTTKGRVPYPGMVNREVLDQVERGYRMPCPPECPESLHDLH} \\ {\tt CQCWR}$

KEPEERPTFEYLQAFLEDYFTST

5

10

20

25

30

35

These residues are incorporated into the appropriate position from the invariant residues listed in Table 4 using the coordinate set provided in Figure 17. Once the template is in place and the catalytic site from pp60^{c-src} has been superimposed onto the template, it is possible to visualize the catalytic site. The site can additionally be refined using the complementary target phosphorylation site for pp60^{c-src}.

RLIEDNEY*TARQGAK

* denotes the site of phosphorylation.

Residues are altered using computer modelling until a fit is achieved for pp60^{c-src} on the template. Thus, residues 184, 166, 172, 220, 208, and 280 from the pp60 c-src sequence have positions in space that maintain those distances disclosed in Table 4. Ionic and hydrophobic amino acid side chains are matched within the catalytic core with complementary residues to create a new inhibitor molecule. Recombinant cAPK is then mutated to duplicate the three-dimensional structure within the core. Crystals of mutated cAPK are analyzed alone or together with a proposed inhibitor. The structure is again analyzed in the context of the invariant residues listed in Table 4.

Positions 52, 72, and 91 are mobile invariant residues whose positions will vary depending on the quality of inhibitor. The distances of these residues are listed in Table 4 for PKI(5-24) and cAPK. It is anticipated that peptide inhibitors of equal affinity for cAPK will have similar distances. Non-peptide inhibitors can be designed that do not produce a rotation, or fraction of fit, exactly in the same direction as peptide inhibitors, such as PKI(5-24). A comparison of the crystal structure of cAPK and cAPK with PKI(5-24) indicate that positions 52, 72 and 91 rotate 12° toward residues 184, 166, 172, 220, 208 and 280. This rotation defines a range of peptide inhibitors. Another strong peptide inhibitor will similarly produce a 12° rotation toward the six residues listed above while residues 52, 72 and 91 may have a smaller angle of rotation for weaker peptide inhibitors.

In addition, there are important points of contact between cAPK and PKI(5-24). The specific contact amino acids on PKI(5-24) are starred and the corresponding points of contact within the catalytic core of cAPK are identified as positions p+1, p-2, p-3, p-6 and p-11. These points of contact are conserved within the catalytic core of all protein kinases

and similar points of contact will be readily identifiable to those of skill in the art for other protein kinases. A sphere of influence having a radius of 11\AA or less, more preferably 6\AA or less, and extending from the inhibitor around the points of contact at positions p+1, p-2, p-3, p-6, and p-11 can be used to define regions that are critical for inhibitor specificity.

5

As described above, in connection with Figure 11, the points of contact can be used to identify the replacements necessary to design appropriate inhibitors or other affectors for a new enzyme. Thus, amino acid replacements are used which form appropriate ionic and hydrophobic interactions at these points of contact. Hydrogen bonding interactions are also preferably used to identify replacements. Of course, the modelling can extend beyond the identified points of contacts in order to provide still further specificity

10

This same type of analysis can be performed with the mutated cAPK that mimics the catalytic core of pp60^{c-src}. Thus, an analysis of the mutant crystals permits one to predict the affinity of a given inhibitor. The inhibitor can be further modified to improve the ionic and hydrophobic interactions surrounding the points of contact using the spheres of influence described above. The angle of rotation of the mobile invariant residues can be used to predict whether or not a given peptide inhibitor will be useful. These changes are all performed within the constraints of the coordinates of Figure 17.

15

A peptide inhibitor that, once modelled has distances similar to Table 4 and meets the design criteria described above can be synthesized and tested for function in vitro or in vivo.

20

25

The coordinates obtained from the binary complex and the resulting template allow us for the first time to fully appreciate the complexity and sophistication of the process by which a protein kinase recognizes its protein substrate. While peptide analogues provide important clues, the diversity of the peptide binding sites and their dispersion over such a wide area on the enzyme surface makes it imperative to have structural data on complexes of the enzyme with affector molecules. The structure of the binary complex of cAPK with PKI(5-24) provides, for the first time, a molecular basis for the rational design of affector molecules, both peptide and nonpeptide, that can target specific protein kinases. Furthermore, because the basic catalytic core of this enzyme is so conserved in all protein kinases, a template based on the crystal structure can also serve as a mold for modelling for other protein kinases.

30

Although this invention has been described using protein kinases as a model system, with cAPK being shown as a specific example of an enzyme for determining the template, the present invention is not intended to be limited to this model. Other changes to the

-54-

methods described herein will suggest themselves to those of ordinary skill in the art. Accordingly, the spirit and scope of the present invention is to be determined with reference to the appendant claims.

WE CLAIM:

1. A method of designing a highly specific affector which exerts an effect on the activity of a first enzyme, said first enzyme being a member of a class of enzymes having a conserved catalytic core, comprising:

5

identifying a second enzyme that is a member of said class in which a first affector can affect the activity of said second enzyme;

forming a first complex of said first affector and said second enzyme;

obtaining data regarding the conformation of said second enzyme at sites greater than 5 Å from the site of catalysis of said second enzyme in said first complex;

10

designing an affector which induces a conformation on said first enzyme at sites greater than 5 Å from the site of catalysis thereof which is homologous to the conformation of the conformation of said second enzyme at homologous sites in said first complex, when said affector is formed as a second complex with said first enzyme; and

15

producing said affector.

2. The method of Claim 1, additionally comprising crystallizing said first complex and obtaining X-ray crystallography data therefrom.

The method of Claim 1, wherein the designing step comprises:

identifying a potential affector likely to induce a conformation on the catalytic core of said first enzyme which is homologous to the conformation of the catalytic core of said second enzyme in said first complex, when said affector is formed as a second complex with said first enzyme; and

25

20

determining whether said potential affector induces said conformation through the use of a technique selected from the group consisting of computer modelling, small angle neutron scattering, and circular dichroism.

- 4. The method of Claim 3, wherein said potential affector comprises a peptide.
- 5. The method of Claim 3, wherein said potential affector comprises a molecule selected from the group consisting of nucleotides, polynucleotides, alcohols, polymers, lipids carbohydrates, sugars, native or synthetic molecules, protein and organic compounds and combinations thereof.

30

6. The method of Claim 1, wherein all of the members of said class have related functions.

PCT/US92/06137

-56-

- 7. The method of Claim 1, wherein the catalytic cores of all of the members of said class have conserved amino acid residues.
- 8. The method of Claim 7, wherein the designing step comprises designing an affector having homologous topography and charge fields that complement the catalytic core of said first enzyme and capable of inducing a conformation wherein the conserved amino acid residues of said first enzyme are in homologous locations to said second enzyme in said first complex.
 - 9. The method of Claim 1, wherein each of the affectors is an inhibitor.
 - 10. The method of Claims 1, wherein each of the affectors is an activator.
- 11. The method of Claim 1, wherein said first affector comprises all or a portion of said first enzyme.
 - 12. The method of Claim 1, wherein said first complex is a holoenzyme.
 - 13. The affector produced by the method of Claim 1.
- 14. The method of Claim 7, wherein said class of enzymes comprises protein kinases.
 - 15. The method of Claim 14, wherein said second enzyme is a viral oncogene product or a cellular homologue thereof.
 - 16. The method of Claim 15, wherein said second enzyme is p60 v-Src from RSV or its cellular homologue, pp60 c-src.
 - The method of Claim 16, wherein said second enzyme comprises cAMPdependent protein kinase.
 - 18. The method of Claim 2, wherein said second enzyme comprises a native mammalian protein kinase.
 - 19. The method of Claim 2, wherein said second enzyme comprises recombinant protein kinase.
 - A method of designing a highly specific affector which exerts an effect on the 20. activity of a first enzyme, said first enzyme being a member of a class of enzymes having conserved residues at an affector binding site, comprising:

identifying a second enzyme that is a member of said class in which a first affector can affect the activity of said second enzyme, said first affector having a dissociation constant with said second enzyme of less than 1 µM;

forming a first complex of said first affector and said second enzyme; obtaining data regarding the conformation of the affector binding site of said second enzyme in said first complex;

10

5

20

25

15

designing an affector which induces a conformation on the affector binding site of said first enzyme which is homologous to the conformation of the affector binding site of said second enzyme in said first complex, when said affector is formed as a second complex with said first enzyme; and

producing said affector.

- 21. The method of Claim 20, wherein said class of enzymes have a nucleotide binding site and each of said affectors is capable of binding to said nucleotide binding site.
- 22. A method of designing a highly specific affector which exerts an effect on the activity of a first enzyme, said first enzyme being a member of a class of enzymes having a conserved catalytic core, comprising:

identifying a second enzyme that is a member of said class in which a first affector can affect the activity of said second enzyme;

forming a first complex of said first affector and said second enzyme, said first complex having at least three points of contact between said first affector and second enzyme;

obtaining data regarding the conformation of the catalytic core of said second enzyme in said first complex;

designing an affector which induces a conformation on the catalytic core of said first enzyme which is homologous to the conformation of the catalytic core of said second enzyme in said first complex, when said affector is formed as a second complex with said first enzyme; and

producing said affector

- 23. A crystallized protein kinase/affector complex having stable decay characteristics over 15 minutes.
- 24. A crystallized protein kinase/affector complex having a Bragg spacing diffraction limit of less than 4Å.
 - 25. The crystallized protein kinase of Claim 24 having stable decay characteristics over 15 minutes.
 - 26. A crystallized complex of a cAMP-dependent protein kinase and an inhibitor thereof.
 - 27. Use of the crystallized complex of Claim 26 in an X-ray crystallography procedure to produce data regarding the three dimensional structure of said cAMP-dependent protein kinase in said complex.

10

5

15

20

25

WO 93/02209

- 28. Use of the data produced by Claim 27 for designing an inhibitor of a second protein kinase which will induce a similar three dimensional structure on the active site of said second protein kinase as the three dimensional structure of said cAMP-dependent protein kinase in said complex.
 - 29. An inhibitor designed by Claim 28.
- 30. A method of preparing a highly specific affector of a first enzyme, said first enzyme being a member of a class of enzymes having a conserved catalytic core with a plurality of invariant residues, comprising:
 - a. identifying a second enzyme that is a member of said class and having a known affector thereof;
 - b. forming a first complex of said second enzyme and said known affector;
 - c. obtaining data regarding the three dimensional coordinates of said invariant residues in said first complex, said coordinates forming a template;
 - d. generating a model wherein said first enzyme is in a conformation in which said invariant residues are in substantially the same conformation as in said template;
 - e. identifying a change in the variable residues in the catalytic core of said first enzyme in the conformation of step (d) when compared to the variable residues in the catalytic core of said second enzyme in the conformation of step (b);
 - f. preparing a modified form of said second enzyme, wherein the modified second enzyme includes the non-conserved change identified in step (e);
 - g. designing an affector which computer modelling indicates will induce a conformation on the catalytic core of said first enzyme in which said invariant residues are in substantially the same coordinates as the coordinates of said template, when said first enzyme is formed as a second complex with the affector designed in this step; and
 - h. producing said affector.
- 31. The method of Claim 30 wherein said change is a non-conserved change in the variable residues.
 - 32. The method of Claim 30, additionally comprising:
 - i. forming a third complex of said modified second enzyme and an affector capable of binding thereto;

10

5

15

20

25

10

15

20

25

30

- j. obtaining data regarding the three dimensional coordinates of the invariant residues in said third complex; and
- k. using the data obtained in step (i) to design an affector which computer modelling indicates will induce a conformation on the catalytic core of said first enzyme in which said invariant residues are closer to the coordinates of said template than the conformation induced by the affector designed in step (g), when said first enzyme is formed as a fourth complex with the affector designed in this step.
- 33. The method of Claim 32 wherein the affector of step (i) is the known affector of step (a).
- 34. The method of Claim 32, additionally comprising modifying the computer modelling used in step (g) in light of the data of step (j), prior to performing step (k).
- 35. The method of Claim 30, additionally comprising obtaining amino acid sequence data relating to the catalytic cores of the first and second enzymes.
- 36. The method of Claim 30 wherein step (f) comprises site directed mutagenesis of a recombinantly produced second enzyme.
- 37. The method of Claim 30, wherein the coordinates of said template are substantially as shown in Figure 17.
 - 38. The method of Claim 30, wherein each of the affectors is an inhibitor.
- 39. The method of Claim 30, wherein said template includes coordinates separated by the distances substantially as shown in Table 4.
 - 40. An affector prepared by the method of Claim 39.
 - 41. A pharmaceutical composition comprising the affector of Claim 40.
 - 42. A method of designing a specific inhibitor for a protein kinase, comprising: obtaining data regarding the three-dimensional structure of a first protein kinase;

using said data in the design of an inhibitor for a second, different, protein kinase; and

producing said inhibitor.

- 43. The method of Claim 42 wherein said first protein kinase is cAMP dependent protein kinase or an analogue thereof.
- 44. The method of Claim 43, wherein the obtaining step comprises crystallizing a complex of cAMP dependent protein kinase and an inhibitor thereof.

PCT/US92/06137

- 45. The method of Claim 44, wherein the obtaining step additionally comprises obtaining X-ray crystallography data on the crystallized complex obtained from the crystallizing step.
- 46. The method of Claim 44, additionally comprising obtaining information regarding the structure of the crystallized complex obtained in the crystallizing step through circular dichroism, small angle neutron scattering, or other chemical procedures.
- 47. Use of the data of Figure 17 or of Table 4 in the design of an affector for a protein kinase.
- 48. A method of preparing a highly specific inhibitor of a first enzyme, said first enzyme being a member of a class of enzymes having a conserved catalytic core with a plurality of invariant residues, comprising:
 - a. identifying a second enzyme that is a member of said class and having a known first inhibitor thereof;
 - b. forming a first complex of said second enzyme and said first inhibitor;
 - c. obtaining data regarding the three dimensional coordinates of said invariant residues in said first complex;
 - d. designing a second inhibitor which computer modelling indicates will induce a conformation on the catalytic core of said first enzyme in which said invariant residues are in substantially the same coordinates as the coordinates obtained in step (c), when said first enzyme is formed as a second complex with said second inhibitor;
 - e. preparing said second inhibitor;
 - f. forming a third complex of said second inhibitor and a third enzyme complexable therewith, said third enzyme having a plurality of said invariant residues;
 - g. obtaining data regarding the three dimensional coordinates of said invariant residues in said third complex;
 - h. using the data obtained from step (g) to design a third inhibitor which computer modelling indicates will induce a conformation on the catalytic core of said first enzyme closer to that in which said invariant residues are in substantially the same coordinates as the coordinates obtained in step (c) than induced by the second inhibitor, when said first enzyme is formed as a fourth complex with said third inhibitor; and
 - i. producing said third inhibitor.

15

10

5

WO 93/02209

20

25

- 49. The method of Claim 48, wherein said first inhibitor is an inhibitory domain of said second enzyme.
- 50. The method of Claim 48, wherein said third enzyme comprises at least 5 invariant residues.
- 51. The method of Claim 48, wherein said third enzyme is a naturally occurring enzyme.
 - 52. The method of Claim 48, wherein said third enzyme is a mutant enzyme.
- 53. A method of determining the efficacy of a first affector in affecting a first enzyme that is a member of a class of enzymes having a plurality of invariant residues among the members of said class, comprising:

determining the three dimensional coordinates of the invariant residues of a second enzyme in a first conformation wherein said second enzyme is in a complex with a second affector that is a strong affector of said enzyme;

determining the three dimensional coordinates of the invariant residues of said second enzyme in a second conformation wherein said enzyme is in a conformation other than said first conformation;

identifying the mobile invariant residues of said enzyme, said mobile invariant residues being those invariant residues at coordinates substantially different in said first conformation than in said second conformation;

determining the three dimensional coordinates of the mobile invariant residues of said first enzyme when said first enzyme is in a conformation wherein said first enzyme is in a complex with said first affector;

comparing the three dimensional coordinates of the mobile invariant residues of said first enzyme in said conformation with the coordinates of the mobile invariant residues of said enzyme in said first conformation, wherein relatively small differences in the coordinates indicate relatively high efficacy of said first affector.

- 54. The method of Claim 53, wherein the step of determining the coordinates of said first enzyme in said conformation is performed using computer modelling of said conformation.
- 55. The method of Claim 53, wherein the steps of determining the first and second conformations comprise obtaining X-ray crystallographic data of said enzyme.
- 56. The method of Claim 53, wherein said second conformation is a conformation produced by a ternary complex.

15

10

5

20

25

PCT/US92/06137

- 57. The method of Claim 56, wherein said ternary complex comprises a protein kinase, a nucleotide and an affector.
- 58. The method of Claim 53, wherein said second conformation is a conformation produced by said second enzyme not complexed with a ligand.
- 59. The method of Claim 53, wherein said second enzyme is the same enzyme as said first enzyme.
- 60. A method of designing specific inhibitors of a first protein kinase having a template defined by a plurality of invariant residues present in substantially all protein kinases, comprising:

10

5

obtaining data regarding the three dimensional coordinates of the invariant residues of a second protein kinase and of the points of contact between said second protein kinase and a known inhibitor thereof, said coordinates being obtained when said second protein kinase is formed as a complex with said known inhibitor;

generating a model of said first protein complex wherein said template is defined by the positions of said invariant residues in said complex;

examining the amino acid residues present in said first protein kinase at positions corresponding to the points of contact in said complex;

designing an inhibitor of said first protein kinase capable of forming ionic and hydrophobic interactions with said amino acid residues; and producing said inhibitor of said first protein kinase.

20

15

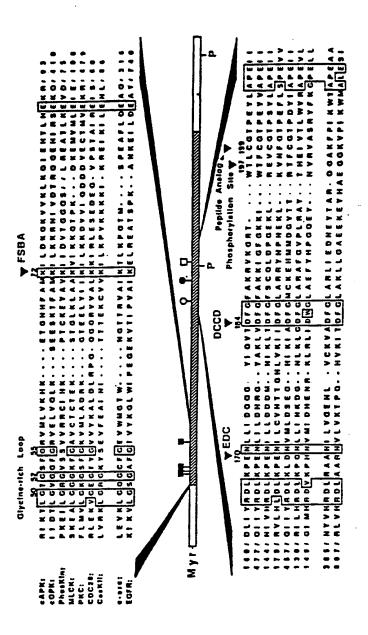
- 61. The method of Claim 60, wherein said second protein kinase is cAMP dependent protein kinase.
 - 62. The method of Claim 61, wherein said known inhibitor is PKI(5-24).
- 63. The method of Claim 62, wherein the points of contact in said complex comprise positions corresponding to the positions selected from the group consisting of P+1, P-2, P-3, P-6 and P-11 along said known inhibitor.
- 64. The method of Claim 60, wherein the positions corresponding to the points of contact in the examining step comprise positions within a sphere having a radius of 11 Å from the coordinates of the point of contact obtained in the obtaining step.

30

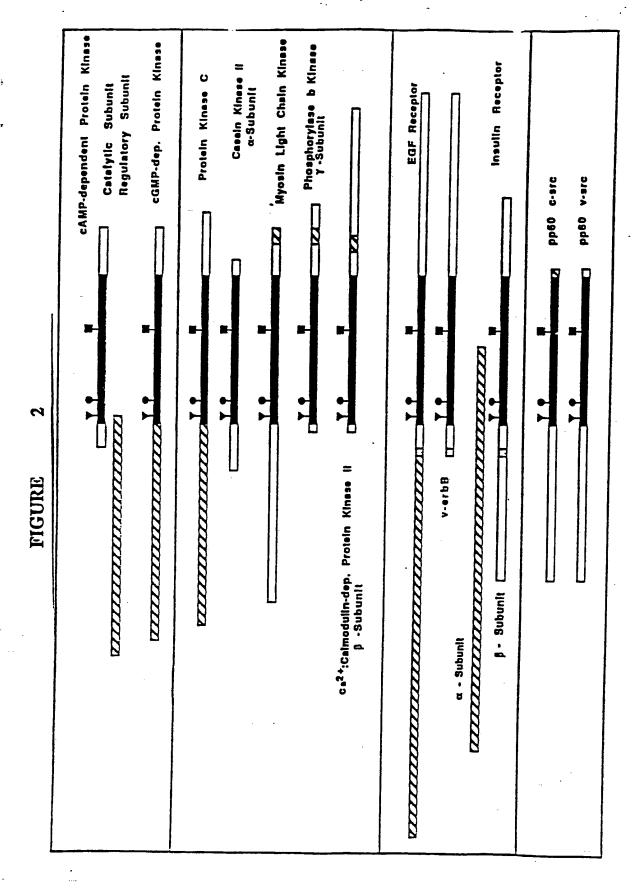
25

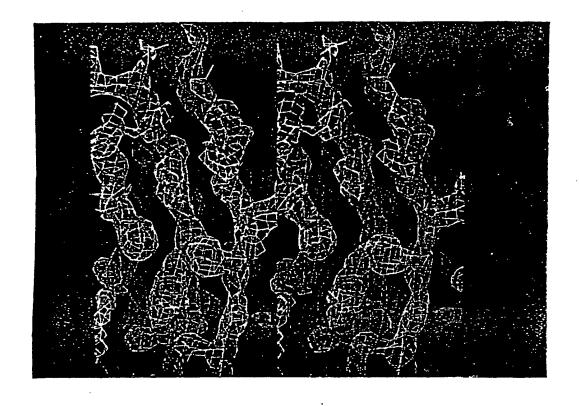
65. The method of Claim 64, wherein the positions corresponding to the points of contact in the examining step comprise positions within a sphere having a radius of 6 Å from the coordinates of the point of contact obtained in the obtaining step.

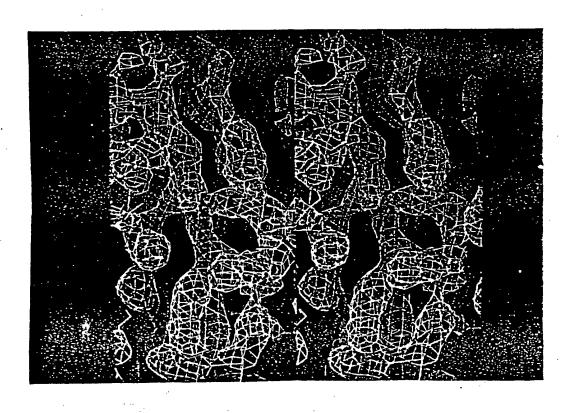
66. The method of Claim 60, wherein the designing step additionally comprises designing said inhibitor to form appropriate hydrogen bonding with said amino acid residues.



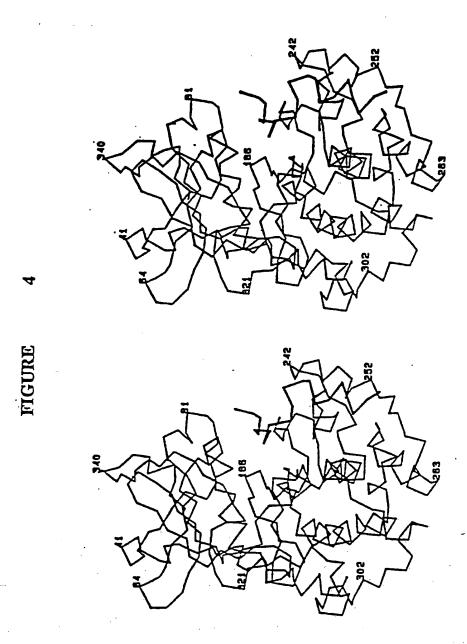
FIGURE

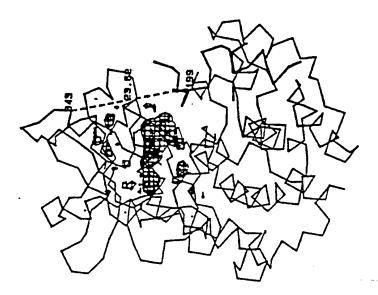


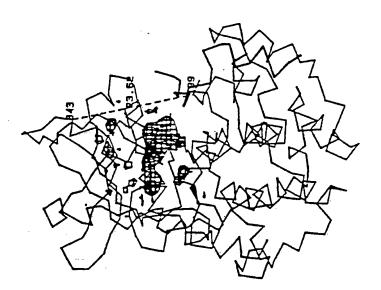




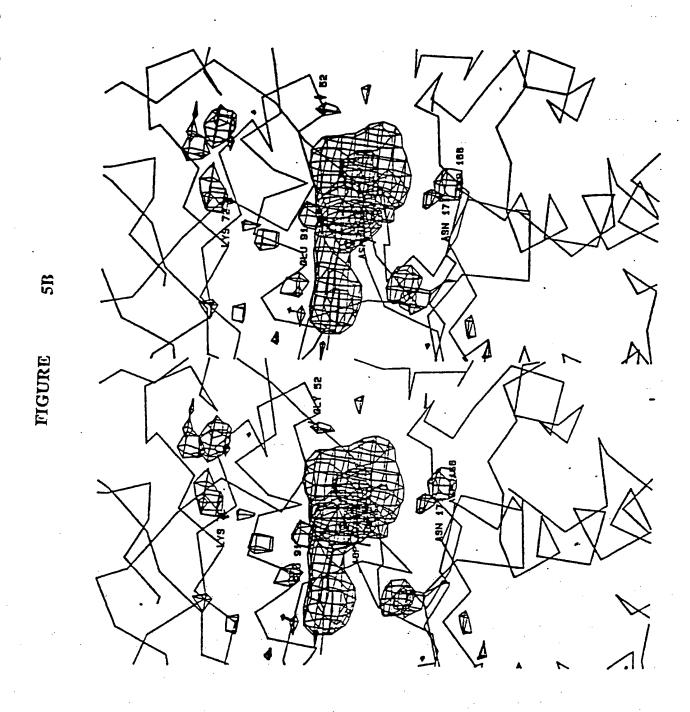
3B

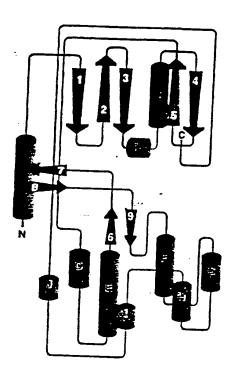






TGURE

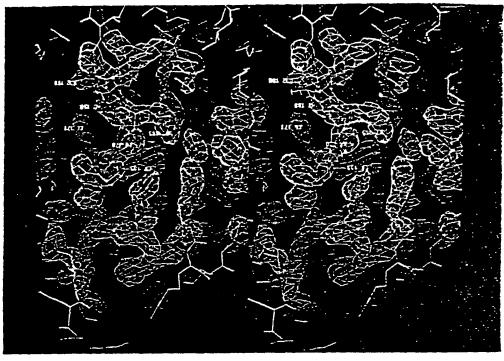


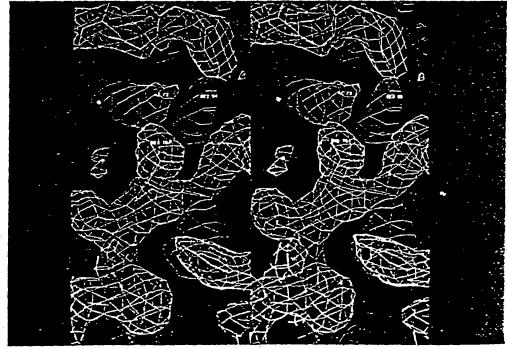


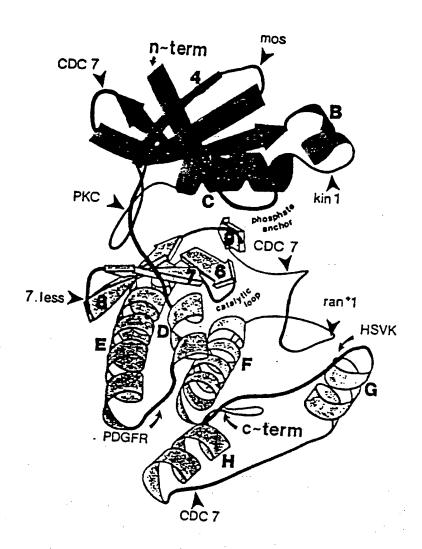
FIGURE

7.13

FIGURE







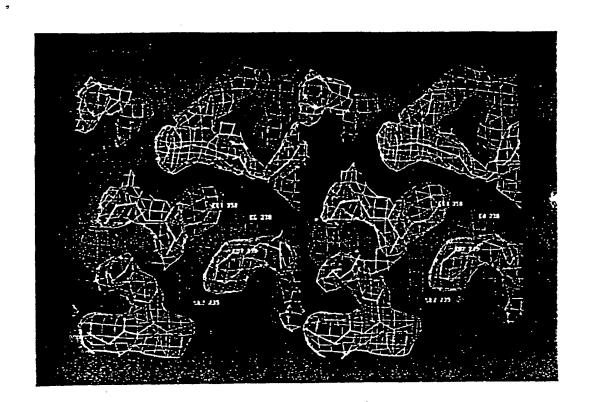
FIGURE

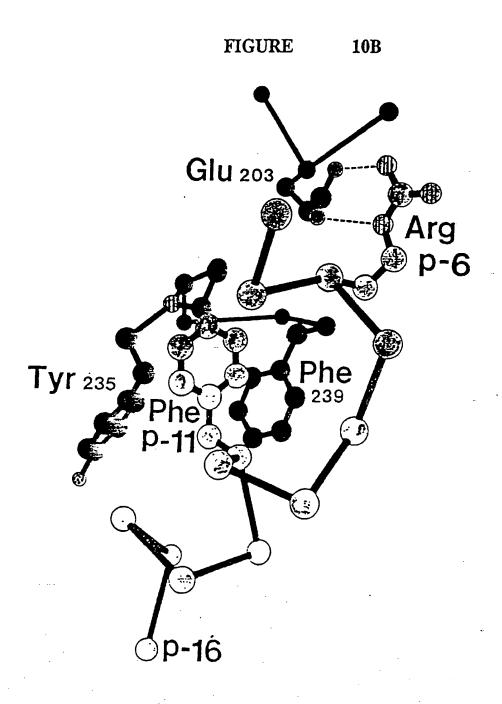
8B

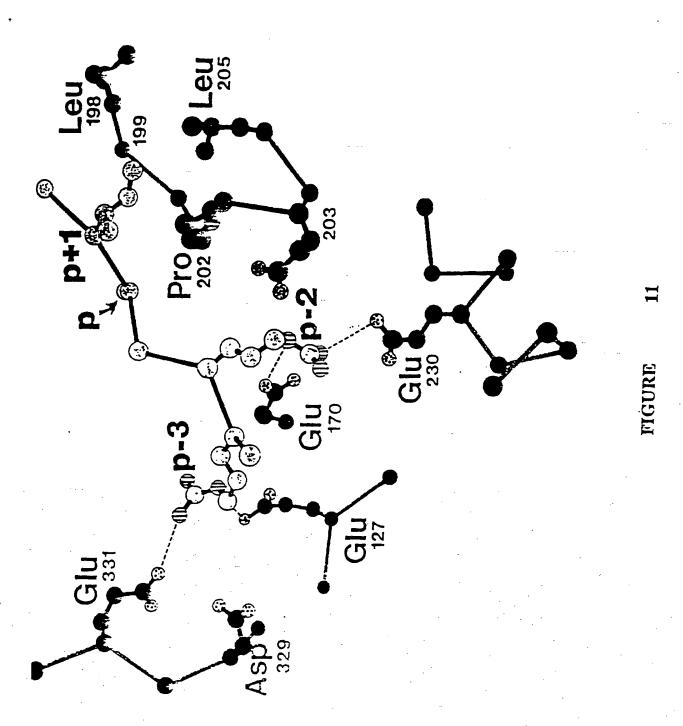
FIGURE

12/49

FIGURE 10A



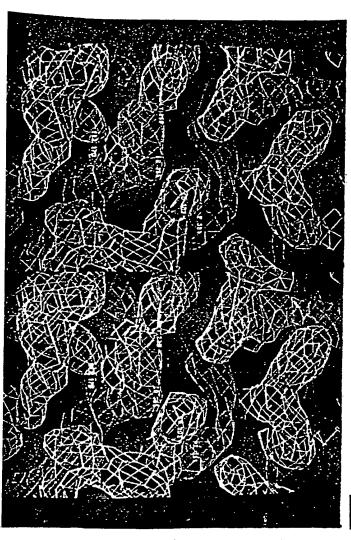


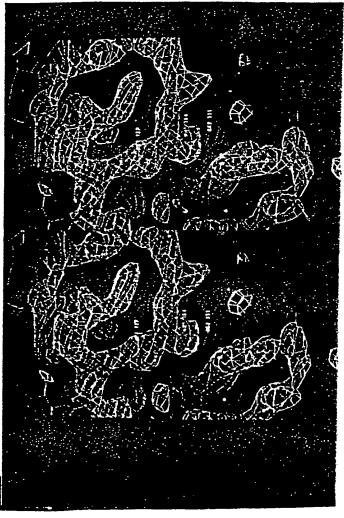


15/49

FIGURE 12B

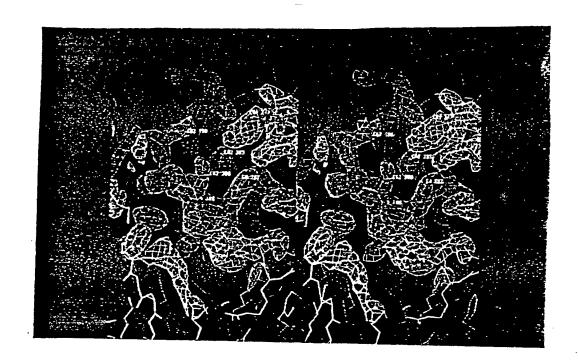
FIGURE 12A





FIGURE

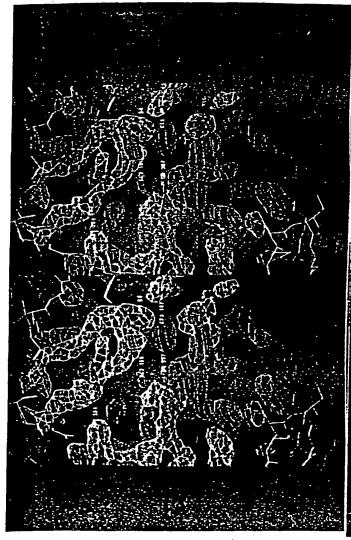
12C

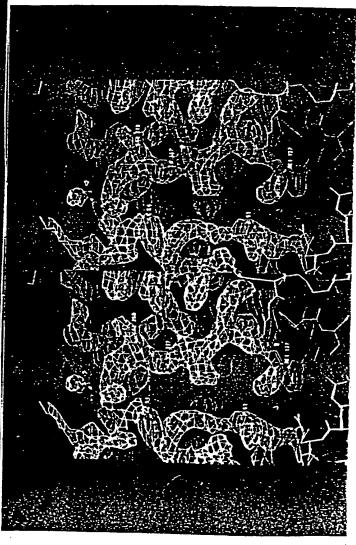


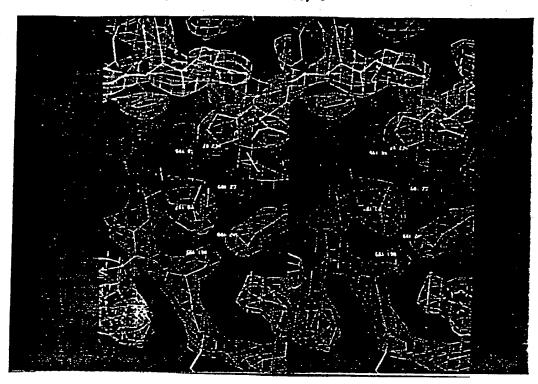
17/49

134

FIGURE

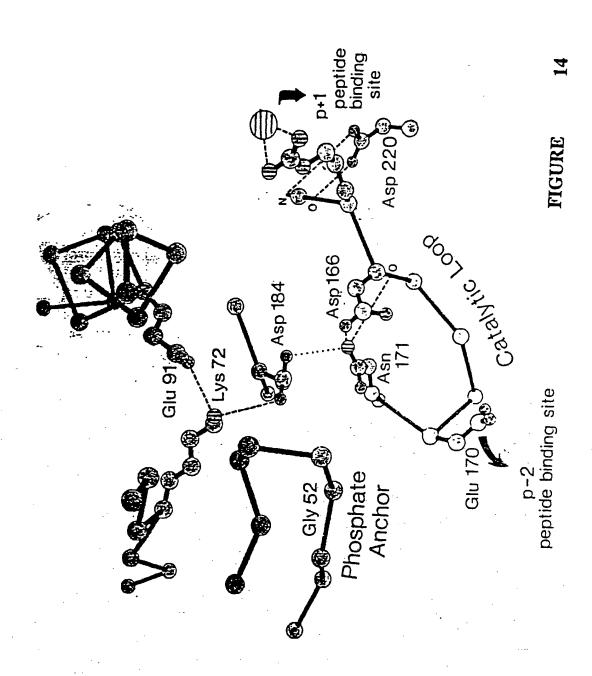


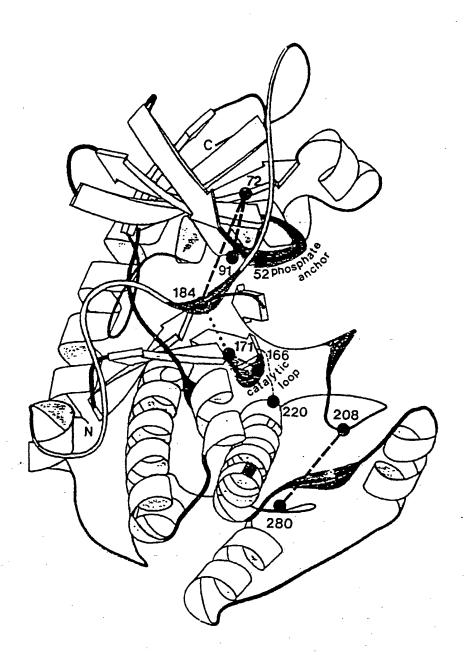




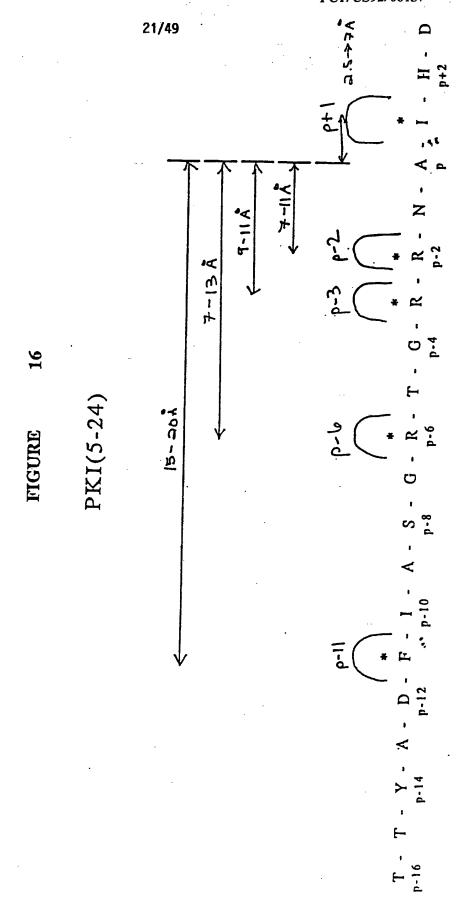
13C

IGURE





FIGURE



10.00 11.00

M104

-
Ξ
\cong
\triangleright
C
Ħ
H

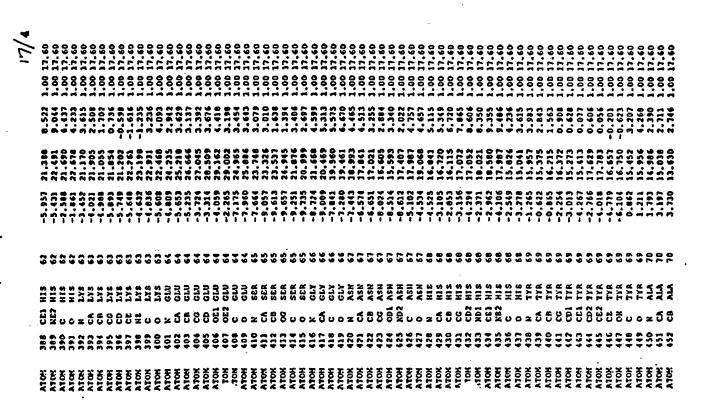
1.00 17	7.00 17	66.	200.4	1.00 17	1.00.1	1.00	7.00	1.00 17	1.00	1.00	00.1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	8	1.00	00.7	2.00	7.00	200.1	00.7	1.00	00	00	1.00 1	000	1.00 1	1 - 17.658 1.00 17.60 1 - 18.988 1.00 17.60	3.00	00.7	00.1	1.00	1.00 1	2	80	1.00	00.0	00.6	
10,934	18.783		12.950	-30.316	-9.710	10.150	-12.285	-13.67	-16.378 -36.378	-17.140	-9.65	-10.208	-9.372	-10.61	-11.406	-30.73	-7.872	2.00	-6.24	-6.403	-3.995	-5.340	-4.320	-2.98	7.42	-6.40	-5.62	-6.49	-6.12	-5.21	-1.43	- (.(67 -5.517 - 1.4.700 -5.021 - 1.4.700 -5.021 - 1.4.700 -5.021 - 1.4.700 -5.021 - 1.4.700 -5.021 - 1.4.700 -5.021 - 1.4.700 - 5.021 - 1	-6.06	5.10		-2.54	-2.72	7.5	0.17	0.62		-0.2	
15	2.	2		15	5 7	<u>.</u>	-	2	:	76	* *	11	2:	``	5		17		: :	2:	::	711C 10	. =	= :	::	6	22	2	2;	2 2	<u>.</u>	AIA 20	2	2 5	200	2	7	≂ ≂	==	π.	1	: 2	:
2 X C	2 5		\$ CG2 VA	۲ د د	š:		10 CB LY	8		2	ه د		5 5	3 8	8	200	U	2	: 5	3	3 8	ê	3 3	t (90	z	ರ ವ	8	ទីខ		٥	< < = 5 € €	8	٥ ,	2	5	5	೭ ಕ	3 2	22	.	o 2	•
TOH.	5	<u> </u>	5	1 <u>0</u> 1	¥ 3	Ž	£0.	10H	3 2	10%		101	¥ §	70.0	A10K	10K	TOH	101	. TO	201	N TO W	HON	ATOH	HOH	100	NIOX	7 0 K	NTOH	ATOH 1	5 *	8	710X	ATOK	ATOK	5 5	ATCH	¥ 20 4	33	15 E	ATOK.	ATOR	ATOM .	

193 196 196	ပဝဍ္	¥	C C C C C C C C C C C C C C C C C C C	-2,481	22.256 23.203 22.080	-9.221	8888	17.60
ijΰυ	: m	53	2 2 2	-1.037	22.231		888	
0 2		ALS GEN	8 8	0.131	24.566	-7.997		
: J (۰.	3	8	0.532	26.157		38	• •
30		3 3	2 0	0.63	27.351		8 6	
2 9	_:	3	80	-0.728	29.137	-6.491	0	17.60
5 2	3 2	5 2	2 0	1907	10.597	5.65	000	
U	1	3	9	0.90	26.284	-4.331	2 8	17.60
0 2		200	5 C	7.0	26.168	-3.429	2.00	17.60
6 3			; ;	2.791	26.451	-4.057	8 8	17.60
ซิ		3	9 :	6.200	27.087	-2.926	3.00	37.60
3 8	. 3		2 2	. 40 . 40 . 40 . 40 . 40 . 40 . 40 . 40	26.933	-2.007	86	17.60
ម	×	23	2	5-196	25.002	-0.934	9.6	
u c		3 :	\$ \$	2.027	26.986	-1.454	1.00	•
×		Š	\$ =	1.719	20.259	-1.672	100	•
5	_	754	Ç	1.042	29.206	-0.839	1.00	
3 5		754 754	;	0.821	30.636	-1.716	1.00	•
8	-	AE P	=	-0.940	31,725	-2.738	000	
8	N	ASP	;	-1.104	29,553	-2.975	2.00	
ه ن		ASP	7 7	-0.233	28.734	-0.247	8 6	•
2		3	÷	-0.607	27.741	0.80	8	
ฮ ซ		3 3	? :	-2.056	27.326	-0.376	1.00	•
38		5 5	7 7	-2.650	25.992	-1.577	9.0	• •
8		3	?	-2,003	26.370	-3.970	1.00	
엉얼	- 1	3 3	? ?	-2.44	27.439	-6.438	80.0	17.60
U	1	3	2	-2.001	26.115	0.529	9	
6 2		5 33	Ç (-3.001	25.511	0.623	00.	
ະ ວົ	_	7 X X	7 3	-0.736	25.786	0.98	5 S	
5	_	¥¥.	Ç	0.476	23.663	1.316	8	
88	=	3 2	35	-0.091	22.666	0-147	2.6	
8	~	3	3	0.030	23.337	-1.126	99.	
ם כ	:: :	2 2	? :	-1.434	21.143	-0.646	1.00	17.60
: 3	٠	1 2	; ;	-1.271	23.585	-2.157	9 5	•
Ü		PHE.	7	0.15\$	25.400	2.984	.00	
0 :		3 5	Ç:	A.028	26.258	2.767	1.00	•
z ī	_	ASP	;;	-0.285	25.003	4.166	80.	
ם כ		7 2 4 V 2 4	::	762.0 ELC 0-	75.57	5,395	99	
0		ASP	=	-0.367	27.20	7.069		•
Б	=	ASP	Į	-0.497	27.262	.308	8	. •
ь	2	ASP	;	0.038	28.246	6.433	00.	
U E		757	Ŧ :	1.042	24.327	5.901	00.	۹.
3 2		ARG	. ¥	2000	25.734	5.83¢	86	17.60
u	5	ARG	*	3,337	23.576	6.654	90	17.60
ü.	<u>.</u>	ARG	\$	4.673	23.942		90.	17.60
ΔI	.	ARG	Ş :	5.607	26.122	6.586	1.00	17.60
6	_	Zec	\$	6.833	24.063	7.474	9.0	٠.

MANY 255 CE PHE 54 17-655 11.039 -4-151 10.01 10.05
ATOM 255 CE PHE 54 17-651 11.039 -4-151 10.01 10.05
ATOM 255 CE PHE 54 15-67 11.039 -4-151 10.01 10.05
ATOM 255 CE PHE 54 15-67 10.05
ATOM 255 CE PHE 55 CE PHE 55

MUCH 258 NE ARG 45 - 1.18 24.601 7.476 1.00 17.60 MUCH 258 NE ARG 45 - 1.18 24.001 17.00 MUCH 258 NE ARG 45 - 1.18 2.738 8.473 1.00 17.60 MUCH 256 NE ARG 45 - 1.18 2.738 8.433 1.00 17.60 MUCH 256 NE ARG 45 - 1.18 2.738 8.433 1.00 17.60 MUCH 256 NE ARG 45 - 1.18 2.738 8.433 1.00 17.60 MUCH 256 NE ARG 45 - 1.18 2.738 8.433 1.00 17.60 MUCH 256 NE ARG 45 - 1.18 46 1.18 2.18 2.18 1.00 17.60 MUCH 256 CD 11E 46 1.18 2.213 20.239 1.28 1.00 17.60 MUCH 256 CD 11E 46 1.18 2.213 20.239 1.28 1.28 1.00 17.60 MUCH 271 0 1.18 46 1.18 2.213 20.239 1.28 1.28 1.00 17.60 MUCH 271 0 1.18 46 1.18 46 1.18 2.213 20.239 1.28 1.28 1.00 17.60 MUCH 271 0 1.18 46 1.18 46 1.18 2.2

A1000 45	, (§ :	2 5	5 B. "5	15,795		3.00 17.
) ;	Ş	2 ;	3.437	õ		1.00 17.
£	. J	Ä	: 5	5.139	77.	1.586	1.00 17.
•	P	TET.	7	5.202		0.40	
•		4	7	4.232	39.775	-0.254	1.00 17.
MICH 46			2 ;	4.673		-0.425	3.00 17.
ATON 46	0	Ā	: 2	6.566		0.105	
¥ ;	0:	Į.	21	7.301	16.966	1.121	1.00 17.6
: =	. đ	EXS	; ;	417		1.052	
NTOH 46	5	173	72	6.639		-2.388	•
¥;	8	LYB	27	9.921		-2.247	-
\$ ¥	8	273	2;	9.686		-3.350	_
7		3 3	2 2	10.493		-1.3	3.00 17.
•		LYS	: 2	9.038	5	-1.932	
+	_	LYE	72	8.510		-2,672	1.00 17.6
▼ •	~ (11.	2	30.064		٠	3.00 17.
• •		118	2	30.516		•	7.00
ATOX C	- 10	9 1	נ ב	10.552		•	1.00
•	ی ،	11	2 2	100		•	1.00 17.
•	~	31.6	5	9.227			200.1
•		37.	5	11.921		٠	1.00 17.
•	.	37	2	32.721		•	1.00 17
	٠.		* ;	12,139		•	1.00 17
-			: :	13.627		•	1.00 17.
TON 48			7	13.044		• •	200
+		ree	7	14.154		•	-
? :		2	Σ;	11.732		•	1.00 17.
-		2	~ ;	14.024		•	1.00 17.
9) 2 > 0	724	<u> </u>	15.371	20.722	' '	~ .
_		AGF	35	15.864		•	1 00 17
_		AST	75	17,023		٠	1.00 17
_		ASP	\$	17.647		•	1.00 17.
100		ASP.	5	18.303		•	1.00 17.
		754	¢	16.046		•	1.00 17.
•		ASP	2 %	17.314		. ,	
170H 496		LYS	92	15.012		•	80.
101		LYS	2	16.296		•	1.00 17.
		27.	۶ ۲	15.825		•	1.00
NTOH SO		LYS	9.	13.657		' 7	1.00 17.
*		LYS	92	14.099		ī	1.00 17
		1.75	36	13,055		7	1.00 17.
.			92	17.820	2	-8.522	1.00 17
^ •			2 :	16.241		-9.139	1.00 17.
		ž 2	2 :	18.665	22.716	-7.856	1.00 17.
			: ;	20.134	•	٠,	37.6
			: =	20.00	36 034	3	2
S			: ::	20,651	26.617	•	••
•			T.	21.571	26.605	: ~:	17.
S			נו	20.096	27.405	.5	
.			נ	20.603	23.534	-7.396	1.00 17.60
· •			Ξ,	21.516	20,663		17.6
10% 514			۶;		21,139	-6.137	17.6
			= ;	21.231	19.971	-5.634	17.6
	3 8	213	2 ;	20.955	19.632	1.75	1.00 17.60
•	3	\$	•	19.67	19.027	-3.51	00 17.6



ره																																																													
17	- 5	2	37.60	17.60	17.60	2		20.7		20.75	9.75	2	20.5	20.5	2		27.60	17.60	37.60	17.60	17.60	17.60	17.60	17.60	17.60	17.60	9.5			200	9	9	17.60	17.60	17.60	9.2	9	2 5		20.0	2.60	7.60	2.60	7.60	7.60	7.60	7.60	9:	9.5	2.7		209	7.60	7.60	2.60	7.60	7.60	9.5	2.6	9.7	7.60
	2	Ş	5 :	3 :	Ē	į	į	į	3 2	į	5 2	ě	į è	Š	Š	Š	Ĉ	8	8	8	8	8	S	8	8	8 8	3	3 8	8	8	8	\$	8	8	8	88	38	38	3 6	3 8	8	2	8	8	8	8	2	2	2	3 2	2	2	9	8	2	8	2	2	7 00 7	200	1.00.1
	136			1			9	6		:	-	5	5	7,5	-	613	922	333	792	261	565	306	50.		262	7			200	3	7	3	413	792	2			23	7	555	136	222	398	261	201	50	2	?;	3 5	18.48.	2	-8.458	-6.439	-7.403	-5.909	-5.348	-5.139	-5.637	-5,714		-2.677
	10.733	988	200			70.0	2.467	6.569	5.756	7.963	7. 668	8 70	8.785	9-502	8.546	9,364	7.682	9.463	9.036	10.384	10,698	12.220	12,334	12.173	16.33		700	90	8.133	7.900	6.797	6.090	6,3(3)	2.01	4.775	0.000		6.199	3.977	3.037	4.387	4.391	5.228	3,500	2.343	. 531	1.959	0.599	653	-2.350	2.337	-3,687	-3.687	-4.361	0.719	1.746	-0.211	-1.436	10.00	777	0.06
	5.798	5.021	7.8.5		N . 2 . 2	6.456	5.085	7.323	8.580	4.368	3.251	4.685	3,739	4.372	5.054	3.654	4.066	2.405	1.377	2.202	0.971	790.1	200.7	1.463		986	-0.06	0.944	0.236	1.120	-0.367	-1.067	0.532	0.224	1.547		-1.059	-1.586	-1.625	-2.740	-4.052	-5.295	-6.191	-5.583	-2.924	-3.281		-2.79	1917	-2.400	-1,075	-2,665	-1.366	-2.133	-3.036	-2.635	-3.657	100.4			-2.890
•	93	6	, F	5	ā	č	ī	×	2	ž	7	\$	56	95	56	95	95	50	5 6		9 1	۵ <u>د</u>) u			96	22	97	9.7	97	97	9 6	D (D 6	* *		8	00	6	۵ ۵	66	66	о. Ф.	2	D (00	00	100	100	300	000	300	100	60,	70.	101		101
-	ARG	ARG	A RG	TIT	11.0	31.0	316	11.0	315	3116	11.0	150	120	LEO	1.50	ng.	1.60	1.20	TEG	¥ :	¥ 2	5 2	5 5	2 2	3	3	3	ž	YIY.	Ą	Ş	414	אר. א	7	5	7	Y.	VAL	ASH	AGN	NSK.	ASK	YSY	ASK	200	725	2 2 2 2	7 7 7 7	2K4	몵	PHC	JH.	PHE	PHB	PHE	Z X	0 6		2 2	2	8
	XH2	υ	0	2				_	_							_	~							_											_	9							001							8					Ç,					و	
	839	6 7 9	650	651	652	653	654	655	656	657	658	689	660	99	662	663	199	665	9 2 9			200	2 5	67.9	613	674	675	919	677	676	619	9	3	700	4 4 5	9	989	687	919	619	9	69	695				6.0	8	669	00 C	9	102	203	č	202	907		9 0	200	5	712
	ATOK	X10X	A10%	ATOK	NTOP.	ATCE	NOT	MOH	10.1	TOT	ATOK	TOK	NTOH	NOH	MG.	101	Ę	2	¥ 5	?		1	2	NO.	ATCH	¥Q.	ATCH	A TOPE	YOX	A TOX	ATOH	HOH	502			TCH	A TOPE	ATOX	ATOK	LTOK	TOX	104	*	Ę				Ž	XOT	TOK	TOK.	\TOK	10%	101	0				8	20%	10%

7			5	17.	7.6	- :	; ;	7	17.	-	-	2:	: :	2	1		Ξ:	Ċ	: :	2	5	Ξ:	:::	7	Ξ.	<u>:</u> :	2:	2	22		17		: :	2	Ç:	: :	5			7.	<u>.</u> :	: :	7.	7.	<u>-</u>	::	::	-		17.6
00.0																• •					•		-		-		-				_		_	_			_			_	~ -	-	_	_	-		-		•	00.
-4,462	2. AD2	-2.40	-5.106	-4,756	-3.409	10.01 01.01	-6.820	-4.771	-5.121	-5.428	-6.683	19.00	-3.92	-2.862	-3.930	-2.692	-2.662	20 674	-2.667	-3,666	-1.778	90.7	-2.339	-3.236	-2.077	-0.627	-0.562	0.547	0.175	0.925	0.107	782	1.615	3.036	5.215	6.722	7.020	7.756	9.069	•	10.677	5.034	4.912	8-498	4.595	2.00	3.140 6.83	7.156	7.953	9.336
22.102	22,393	23.503	22,933	24.046	24.335	18.602	16.601	17.664	16.609	15,356	35.405	14.507	16.398	16.236	16,364	16.318	17.672	18.803	15.004	14.620	14.075	12.730	30.914	11.751	11.781	12.656	11.663	11.401	9.611	11.008	11.176	11.781	10.775	11.496	12.358	12.236	12.114	12.189	12.010	13.663	11.70	9.490	1	7.622	6.558	5.136 6.55	7.786	7.895	7.888	6.111
6.304	7.223	7.946	7.549	6.270	4.284	6.281	5.877	5,943	4.982	5.721	7 454	8,614	4,069	1.621	2.755	2.041	160.4	740	1.319	0.445	3.636	2.097	2.963	4.298	S. 646	6.026	-0.836	-1.762	-6.029	-5.01	-5.961	-0.833	0.339	-1.22	-0.347	-0.007	1.336	-0.914	-0,515	0.873	-0.78	-1.952	0.223	0.03	0.071	167.0	0.668	1.030	-0.195	0.253
711																																121																		124
art art	~	TYR			17 E				1 1 1						λ.																	3 3												74				VAL		3
2 3	G	3	200	נַ בַּ	₽	Ų	0 :	k (5 8	3 2	9	충	U	0 :	z i	5 5	5	C 95	Ų	0 :	E (5	8	ŝ	٤ د	٥	×	ฮ ฮ	8	8	062	້	0 :	≥ ਹੈ	6	8	5	200	2	5 8	ن	•	*	5 5		3		0	×	5
170	045	970	?	7 4	650	653	652	M 4	\$ 0 P	¥ 5 6	657	858	829	960	198	7.0	96	965	990	653		670	471	672	874	\$7 \$	948	# C	979	2		683	684	0 0 0 0 0 0 0 0 0	667			091	268		638	968	687			606	805		906	
ATOK	ATOK	ATOK	A 10.5	šä	104	š	¥ ;	£ 2	A705	Ŧ	NTOH.	F10H	HO!	20	A 100	X 01.0	10	ğ	TOX	* 8	ATOX	ž	₹	z 2	5 Z	ATOH	ATOH	Ę X	NTOH	NTOH.	Ç Z	ğ	ğ	××	ATOK	¥ :	100	70%	¥ ;	100	10%	30%	¥ :	F 3	A DE	. T.	HOLL	F-0	z :	2 2



29/49

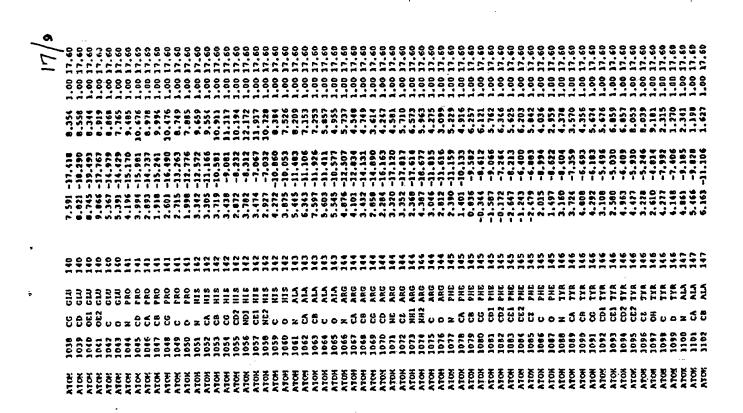
1,00	0	1.00 17	7.00 17	1.00 17	7.00.1	1.00 17	1.00 17	1.00 17	1.00 17	9	7.00.1	1.00 17.6	1.00 17.6	7.00 17	3.00 37.6	71.00	1.00 17.	1.00 17.	7.00 37.	1001	1.00 17.	1.00 17.	1.00 17.		90	1.00 17.	1.00 17.	1.00 17.	14.691 1.00 17.60	1.00 17.	2.00	9	1.00 17.	1.00 17.	1.00 17.	1.00 17.	1.00 17.	7.00.	1.00 17.	1.00 17.	1.00	1.00 17	1.00 17.	1.00 17.	1.00 17.	1.00 17.	1.00 17.	1.00	1.00 17	1.00 17.6	1.00 17	1.00 17.6
-1.154	-2.149	-1.999	-0.934	-2.931	762.1-	-1.127	-1.223	-0.173	-0.386	0.00	0.517	-0.425	1.191	-2.576	320.6-	14.549	-4.780	-6.216	-3,902	-5.820 -S. 60	-6.50	-5.442	-6.354	1000	-6.033	-9.123	-10,485	-11.30	11.517 -13.314	-16.630	-15,398	-9.037	-6.319	-9.752	-9.239	-7.736	-6.056	-5.485	-5.014	-4.960	-12.246	-11.054	-13.766	-13.527	-12.396	-17.809	11.533	-13.692	-13.567	-12.756	-14.602	087.51-
133	133	133	133		333	7.	134	7	130	3	136	334	136	336	135	135	135	33	135	138	135	136	967	<u>.</u>	5	137	55	36	121	2:	137	133	137		2 2	136	136	138	136	2	136	139	139	139	139	6		î	139	£ :	2 5) (
		_	-					2 4 C	-	ARG																			A G									32			2 2								SEP		3 5	
-	7. K						_	~ .	3 2	es Kr	70 98	B7 KH		0 C	, –	~	~												NC NC																							
•	ATOH 9	•	•		•	•	•	ATOK	i ä	ATOK 9	6	6	o . (ATON	. •	•	•										ATOK 10		1000								ATOM 10		ATOH 10		A10H			_								ATOK 10

MATCH STORY OF A ALM 1144 1150 1-1519 1-1519 1-1510

WO 93/02209 PCT/US92/06137

30/49

0.154 0.157 1	0.075 -1.057 1.00 17.	0.450 0.545 1	0.681 -0.4(2 3.00 17.	1.192 0.200 1.00 17.	9.329 -1.065 1.00 17	9.209 -2.233 1.00 17	6.272 -0.373 1.00 17 6 956 -1 000 1 00 17	5. 334 O. 0. 1. 00 1. 1. 00 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1.	5.699 0.619 1.00 17	5.216 1.942 1.00 17	4.058 2.027 1.00 17	5.860 3.001 3.00 17	6.493 -1.879 1.00 17	5.808 -2.846 3.00 37	6.760 -1.643 3.00 17	-2.549 1.00 17	6.350 -1.303 1.00 1.	5.239 -0.011 1.00 17	3.784 -1.024 1.00 17	7.261 -3.729 1.00 17	6.744 -4.825 1.00 17	8.571 -3.650 1.00 17	116 -4.841 1.00 J7	51 00 1	520 -3.757 1.00 17	746 -5.800 3.00 37	103 -6.862 3.00 17	144 -5,418 1.00 17 50 -6.321 1.00 17	306 -5.618 3.00 17	119 -4.052 1.00 17	7, 100 1. (27.7) - 600	119 -6.654 1.00 17	70 -8.003 1.00 17	592 -5.961 1.00 37	765 -6.311 1.00 17	-3.574 -4.148 1.00 17.60	136 -5.329 1.00 17	138 -7.366 1.00 17	71 00 1 2 1 2 2 5 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	191 -0.114 1.00 17	7. 565 1.00 17	76 -6.912 1.00 17	110 -6.450 1.00 17	503 -6,161 1,00 17	578 -5.636 3.00 17	27 -5.760 1.00 17	113 -9.473 1.00 17	20 - 20.440 1.00 17	156 -10.625 1.00 17	10 -10.233 1.00 17	78 -11.350 1.00 17	.340 -10.903 1.00 17	1.496 -11.347 1.00 17	17 00 T 367 70 T 367 6
.386 -	7	153 -1	. 153	- 926	606	200	177	7	49 -0.091	49 -0.083	456	49 0,347	49 2,564		3,659		7,10		6.139	4.849	5.022	•	4,125	3,460	5.193	3.432	3.867	1.209	-0.145	-0.027	-2.313	1,520	1.196	2.114	2.544	2,112	3.451	3.592	4.501	5.548	6.597	7,655	6.628	6.715	9.674	2007	7.03.	4.013	3, 487	2.494	1.047		0.839	- 117. 2.11
C ALA					•	\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \			3	₹	35 T	35	STS C	3	31:	911	311	11.6	11.6	371	115	7	7 4	ZY.	VAI	YY!	3 :	223	780	LEU	797	183	121	Ĕ i	THE 91-	0C1 THR	THE	THR	PKE	PHE	3 1	7 K	PRC	PAC	PRC	3 5	בים בים	3	3	SE.		3 5		3
ATOR 1103	~		ATC: 1106		-	•	•	_		_	_	N70H 1116		TOTAL TOTAL		ATOM 1121	_	_	_	ATOK 1125	_	_ •	A70H 3129	_	_	70H 1132		-	_	ATOX 1137		_	_			136	_		ATOM 1169			ATOK 1151	_	-		•	-	-	_	_		ATOM 1164		



NY 1331 CC | 115. | 15. | 13.56 | -1.099 -12.721 | 1.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.00 | 17.0

MUCH 1166 O CUM 1155 1.007 - 6.420 - 12.662 1.00 17-65

```
MATCH 1355 K TIK 179 -1.755 -2.405 6.425 1.00 11.60

MATCH 1355 CG TIK 179 -1.755 -2.407 1.00 11.60

MATCH 1355 CG TIK 179 -1.554 -1.307 1.00 11.60

MATCH 1355 CG TIK 179 -1.554 -1.307 1.00 11.60

MATCH 1355 CG TIK 179 -1.554 -1.307 1.00 11.60

MATCH 1355 CG TIK 179 -1.554 -1.307 1.00 11.60

MATCH 1370 CG TIK 179 -1.554 -1.307 1.00 11.60

MATCH 1371 C G TIK 179 -1.554 -1.307 1.00 11.60

MATCH 1372 C G TIK 179 -1.554 -1.307 1.00 11.60

MATCH 1372 C G TIK 179 -1.535 -1.535 1.007 1.00 11.60

MATCH 1373 C G TIK 179 -1.535 -1.535 1.007 1.00 11.60

MATCH 1374 C TIK 179 -1.535 1.008 1.001 1.00 11.60

MATCH 1375 C G TIK 179 -1.535 1.008 1.001 1.00 11.60

MATCH 1375 C G TIK 179 -1.535 1.008 1.001 1.00 11.60

MATCH 1375 C G TIK 170 1.004 1.008 1.008 1.001 1.00 11.60

MATCH 1375 C G TIK 170 1.004 1.008 1.008 1.001 1.00 11.60

MATCH 1375 C G TIK 170 1.004 1.008 1.008 1.001 1.00 11.60

MATCH 1375 C G TIK 170 1.004 1.008 1.001 1.00 11.60

MATCH 1375 C G TIK 180 1.004 1.008 1.001 1.00 11.60

MATCH 1375 C G TIK 181 1.005 1.008 1.001 1.00 11.60

MATCH 1375 C G TIK 181 1.00 1.005 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1.001 1
```

NUMBER 1239 CC 51M 170 16,011 1.552 5.052 1.00 17.60

NUMBER 1239 C 51M 170 11,1315 3.158 4.373 1.00 17.60

NUMBER 1239 C 51M 170 11,1315 3.158 4.373 1.00 17.60

NUMBER 1230 CC ARR 171 10,1316 2.331 4.373 1.00 17.60

NUMBER 1230 CC ARR 171 10,1316 2.331 4.373 1.00 17.60

NUMBER 1230 CC ARR 171 10,1316 2.331 4.373 1.00 17.60

NUMBER 1230 CC ARR 171 10,1316 2.331 4.373 1.00 17.60

NUMBER 1230 CC ARR 171 10,1316 2.331 1.00 17.60

NUMBER 1331 CC ARR 171 10,1316 2.331 1.00 17.60

NUMBER 1331 CC ARR 171 10,1316 2.331 1.00 17.60

NUMBER 1331 CC ARR 171 10,1316 2.331 1.00 17.60

NUMBER 1331 CC ARR 171 10,1316 2.331 1.00 17.60

NUMBER 1331 CC ARR 171 10,1316 2.341 1.00 17.60

NUMBER 1331 CC ARR 171 10,1316 2.341 1.00 17.60

NUMBER 1331 CC ARR 171 10,1316 2.341 1.00 17.60

NUMBER 1331 CC ARR 171 10,131 1.00 17.60

NUMBER 1331 CC ARR 171 10,131 1.00 17.60

NUMBER 1331 CC ARR 171 10,131 1.00 17.60

NUMBER 1331 CC ARR 171 1.00 17.60 1.00 17.60

NUMBER 1331 CC ARR 171 1.00 17.60 1.00 17.60

NUMBER 1331 CC ARR 171 1.00 17.60 1.00 17.60

NUMBER 1331 CC ARR 171 1.00 17.60 1.00 17.60

NUMBER 1331 CC ARR 171 1.00 17.60 1.00 17.60

NUMBER 1332 CC ARR 171 1.00 17.60 1.00 17.60

NUMBER 1332 CC ARR 171 1.00 17.60 1.00 17.60

NUMBER 1332 CC ARR 171 1.00 17.60 1.00 17.60

NUMBER 1332 CC ARR 171 1.00 17.60 1.00 17.60

NUMBER 1332 CC ARR 171 1.00 17.60 1.00 17.60

NUMBER 1332 CC ARR 171 1.00 17.60 1.00 17.60

NUMBER 1332 CC ARR 171 1.00 17.60 1.00 17.60

NUMBER 1332 CC ARR 171 1.00 17.60 1.00 17.60

NUMBER 1332 CC ARR 171 1.00 17.60 1.00 17.60

NUMBER 1332 CC ARR 171 1.00 17.60 1.00 17.60

NUMBER 1332 CC ARR 171 1.00 17.60 1.00 17.60

NUMBER 1332 CC ARR 171 1.00 17.60 1.00 17.60

NUMBER 1332 CC ARR 171 1.00 17.60 1.00 17.60

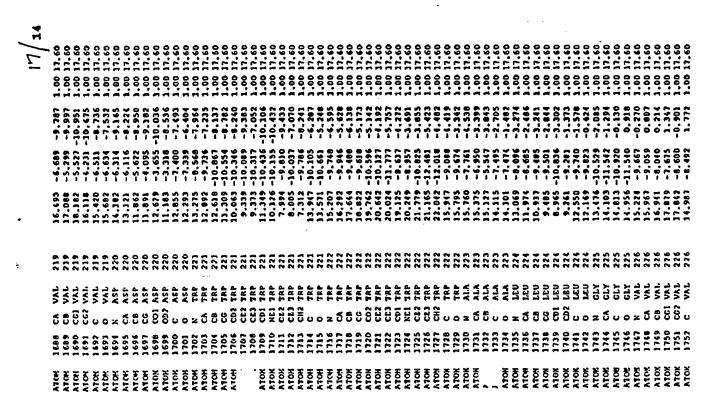
NUMBER 1333 CC ARR 171 1.00 17.60 1.00 17.6

PCT/US92/06137

THE HARRIER COURT OF THE FER T

200	1623	8	312	230	29.662	-5.033	-6.015	00.	
5 2	1626	ع د	1 -	2 5	26.361		,		
ATO.	1626	2	32	217	28.186		2 2		
ž	1627	ฮ	nar	233	29.045		Ξ	1.00	_
ATOK	3628	5	27	211	26.979		-11.168	00.	17.60
5 2	2630	8 8		3 5	30.03			00.0	2 2
₹	163	8	121	211			٦,	1.0	2
¥ 2	1632	y e	2	:::	26.330		Ξ:	1.00	2
	1636) 7		212				9.5	37.60
<u> </u>	1635	: ర	87.8	212			. 7		17.60
704	1636	8	SER	212			7	3.00	2
10	1637	8	SER	212	29,350		-13.763	_	7
ž	-	Ų,	SER	212	27,352		-15.174	_	:
žį	1639	0 2	82 k	212	28.027		-16.132		-
Ę		. ქ	Š	13	26.003		-16.542	• • • •	-
절	•	8	LYS	213	26.000		-16.147	•	7
MOL	1643	8	276	213	27,382		-15.687	_	7.
ě	191	8	2. 2.	22	27.370		-14.613	•••	5
	1646	3 2	2 2	213 F19	76.464		-15.409	•	::
ž	3647	è u	i s	213	24.655		-17.235		::
TOF.	3646	0	LYS	213	24.746		-18,186	•	7.
Į.	3669	z i	CLY	214	23.926	7	-16.893	• • •	Ξ.
	1650	đ .	ב פני	712	22.655	7 1	-17.551	•	<u>.</u> :
Ž	1652	0	ë	21	21.833	7	-16.683		::
ž	1653	×	Ĕ	215	20.614	7	-16.458		1
ž	3651	5	#	235	19.646	ů.	-15.619		Ξ;
ğ	1656	8	Ĭ	215	20.03	77	-11.136		::
100	1657	8	ĭ	215	18,955	7	-13.164		2
10X	1658	8	Z.	215	16.911	7	-12.743	-	7
Ž (1659	õ	ž	235	21.120	7.	-13.960		<u>:</u> :
ATOK	999	20	ž	215	19.003	-1.228	-13.542		בנ
ž	1662	3 ₹	Į,	215	19.756	, 0	-12.505		:::
E S	1663	b	TYR	215	18.269	7	-16.107		=
ž	1664	0:	TY.	235	18.007	7,	-16.966		<u>.</u>
2 2	1669	z 0	ASK	236	17.345	7 7	-15.675		į
₹	1667	8	YSK	216	15.806	7	-16,901		2
2	1660	8	ASK	216	16.398	7	-16,334		Ę.
	1669	8	7 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	326	16.693	7	-15,132		Ċ:
Ę	1673	į U	NSK NSK	216	15,153	3	-14.669		
ž	1672	0	ASR	216	15.643	7	-13,540		5
26	1673	æ	LY\$	717	13.889	٦	-14.94	-	7
ž	1674	5	Ľ.	217	12.052	Ŧ	-13.996		11
ž 2	1675	8	. YS	212	11.612	7	-14.631		2:
<u> </u>	1677	3 8	2 2	217	13.06	7 1	-16.021		:
ATOK.	1678	3	rks rks	712	33,315	7	-17,221		;;
ğ	1679	2	LYS	217	12.779	Ŧ	-17.387		1
ğ	1680	Ų	LYS	217	13,337	7	-12.914		11
ž	1661	: ۵	143	2	12.702	٠,	-11.069		:
100	7991	4 8	? :	2 :	14.363	,	-13.100		::
	1665	5 8	1 7	218	14.77	-7.089	-12,135		7.5
ATOH	1685	ני ל	1	2 2	15.462	7	-10.886		17.6
₹	3686	0	Ę	218	15.568	7	-9.850		2
ATOH	1687	*	X	23.9	15.901	-7.300	-10.952	1.00	7
	,	í	í	:			1	•	•

-8.756 2.945 1.00 17	1.855 1,335 1	-7.404 2.235 1.00 17	-6.862 3.546 3	-6.168 7.472 3.00 17	25.881 1.831 1.00 1.	-8.573 3.039 1.00 1	-8.492 4.255 1.00 17	-9.708 2.468 1.00 17	-10.836 3.225 1.00 1	-12.015	-11 656 1 461 1 00 10	-12.704 0.299 1.00 17	-11.125 4.256 1.00 17	-11.662 5.341 1.00 17	-11.388 4.025 1.00 17	-12.035 4.312 3.00.33	-12.623 5.093 1.00 17	-12.048 6.189 1.00 17	-12.635 6.743 1.00 17	-113.764 4.578 1.00 17 -14.186 5.155 1.00 17	-13.795 6.194 1.00 17	-14.356 6.616 1.00 17	-10.702 6.046 1.00 17	-11.107 7.226 1.00 17	-0.436 6.593 3.00 17	-7.203 5.656 3.00 17	-5.709 6.007 1.00 17	-4.664 6.559 1.00 l7	-4.267 5.813 1.00 17	-6.448 7.669 1.00 17	-0.531 6.865 1.00 17	-8.506 7.203 1.00 17	-8.524 8.073 1.00 17 -8.581 7.217 1.00 13	-7.355 6.413 1.00 17	8.603 -7.416 5.592 1.00 17.60	-0.164 4.278 1.00 17 -0.700 0 020 1 00 13	-9.434 10.207 1.00 17	-10.972 8.621 1.00 17	-12.131 9.50s 1.00 17	-13.168 6,746 1.00 1/	-12.872 11.426 1.00 17	-12.048 9.752 1.00 17	-12.212 10.519 1.00 17	-12.390 9.667 1.00	-11.314 12.321 1.00 17	-9-830 11.08 1.00 17	-0.652 12.007 1.00 17	-7.776 11.438 1.00 17	-6.645 11.864 1.00 17	-8.336 30.415 1.00 17	-7.705 9.737	-9.367 10.61¢	-9.817 12 062
- 942	u	22	£ 5	: ני	727	22	227	228	278	27.	728	228	228	228	229	228	223	229	229	52	523	553	£ 2	230	22	230	230	230	270	230	230	23	3 3	233	23	231	112	232	232	222	22	233	2 :	233	233	2	234	234	234	235	215	235	235
VAL	120	3	3	3	2	2) TEC	31	3 :	3 5	111	1	1	11	É	. X	¥.	7	Ĕ	, <u>7</u>	ţ	TYR	Ĕ	¥ 12	3	CI5	3 :	3 5	3 2	3	275	MET	HEY.	HET	767 5	127	7.27	YIY S	1	1	7	Y	2:	5 5	3	GLY	CILY	פרג	CLY	# ?	- T	17.	TYR
٥																									-										8 5							-											8
1753	1754	1755	1756	1756	1759	1760	1761	1762	1764	1765	1766	1767	1768	1769	27.70	2771	1773	177	3775	77.	1778	1779	1760	1782	1703	1764	1765	1785	1788	1789	1790	1791	1793	1791	1795	1797	1796	1799	1800	1802	1603	1604	1005	1802	1608	1609	1830	1011	1012		1015	1616	1617
ATOH	ATO:	ATO	A 104	704	ATOX	ATOK	ATOK	ATOK	1 0 T	ATOK ATOK	ATOK	ATOK.	ATOK	ATOM	Y LOW	ATOH	ATOH	A DE		ATOH	ATOH	ATOH	ATOH	ATOH	ATOH	ATOH	ATOM	1014	1014 1014	ATOH	ATOR	A.704	ATOK	ATOH	A TOK	ATOK	ATOK	A700	ATOX	ATOX	MTOX	ATOK	ğ	ATOK	ATOK	ATOK	Arok	ATOK	ATOK	ATOK	ATOK	NOK	ATOK

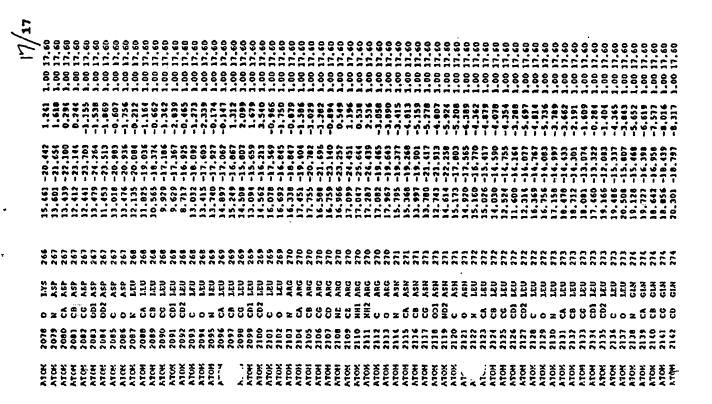


WO 93/02209 PCT/US92/06137

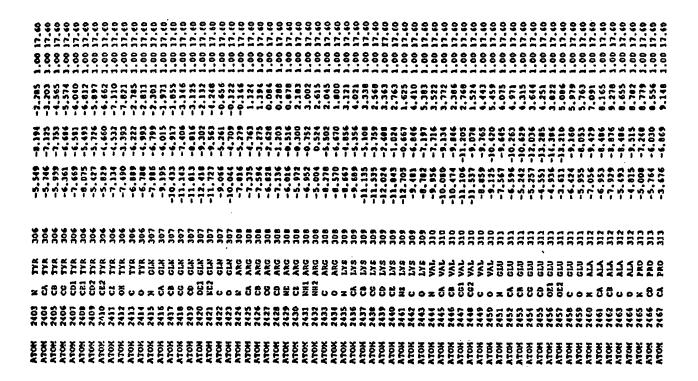
36/49

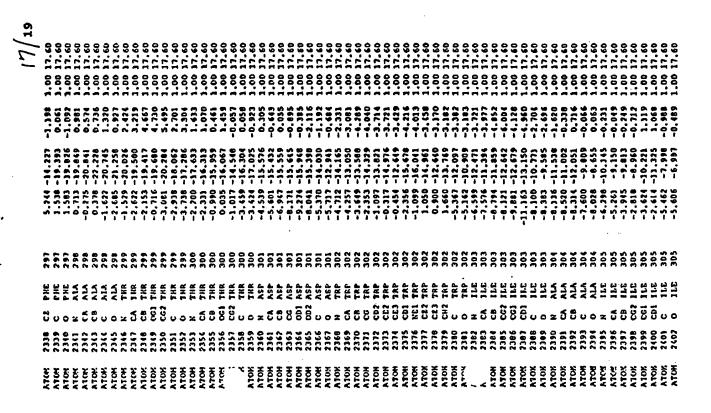
11111111111111111111111111111111111111	-19.617 11.184 3.00 17.	-20.464 10.316 3.00 1	-20.223 32.436 3.00 37.	71.00.E 12.51 1.00	-20.569 15.064 1.00 17.	-23.434 12.986 1.00 17.	-22.451 12.691 1.00 17.	-20.217 13.135 1.00 17.	-10.103 13.736 1.00 17. -18.941 14.251 1.00 17.	-17.562 13.411 1.00 17.	-17.094 13.752 1.00 17.	-16,566 13,455 1.00 17.	-15.511 13.189 1.00 17. -15.864 14.173 1.00 17.	-19.916 12.003 1.00 17.	-19.314 32.000 1.00 17.	-20.427 10.901 1.00 17.	-20,100 9,641 1,00 17.	-18.329 R.566 1.00 17.	-17.432 9.495 1.00 17.	-17.612 7.486 1.00 17.	-16.067 9.347 1.00 17.	-15.576 8.275 1.00 17.	-21.005 9.003 1.00 17.	-20.459 8.362 1.00 17.	-ZZ.3UZ 9.171 1.00 17.	-22.662 4.169 1.00 17.	-21.930 6.986 1.00 17.	-23.265 7.063 1.00 17.	-22,333 6,503 1,00 17, -24,469 6,494 1,00 17,	-24.632 5.191 1.00 37.	-26,302 (.915 1.00 17.	-23.060 (.163 1.00 17.	70 -23,119 3,535 1,00 17,60	-23,700 6,267 1,00 17.	-22,772 3.e5h h.00 17.	-23.860 3.522 1.00 17.	-21.658 3.977 1.00 17. -24.900 2 643 1 00 13	-21.573 3.004 1.00 17.	-21.298 1.617 1.00 17.	-70.600 4.065 1.00 17.	-18.608 5.099 1.00 17.	-17.191 4.414 1.00 17.	-16.531 3.829 1.00 17.	-16.403 6.091 1.00 17	-19.00# 2.354 1.00 12.	-20.506 3.944 1.00 17	-20.685 3.489 1.00 17.	-21.871 4.108 1.00 17	-21,725 5,535 1,00 17.	-23,052 5,806 1,00 17,	-23.329 7.936 1.00 17	1. 00 . 000 C 200 CC
	~	~	ER :	~ ~	. ~	~	~	•		. ~	~	~		. ~	~	•	•	• •	~	~	~ .	• ~	~	~ .		~	~	e4 e	~ ~	~	•••	• ~	~ ~ .	~ ~	~	~ .		~	~ (~ ~	. ~	~	~ •		• 17		~	•				
PRO 25 SER 25 SE	-		-																																											2	_					_
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	ATOK 20		ATOK 20		ATOK 20			ATOM 201					10X 20								ATOM ZO			25							ATOK 20			70X 20			NTOK 20			ATOK 20				TOTAL 20	. ~	~						

16																																																							
1																																																				17.60			
	1.00	00.	90.	1.00	90.	3 6		0	00.	3.00	1.00	9	00.	3 6	00	00.	00.	00.		00.	3.00	1.00	8	3 6	00	9	1.00	86		00.	3.00	8.6	8 8	8	1.00	00	3 6	3 8	1.00	9.0	86	3 2	8	1.00	1.00	9.5	8.6	3 6	1.00	0	1.00	3.00	90.	9 6	80.
	-3.640	-1.462	-1.968	-3.323	-3.663	4.637	517 S-	-6.423	-6,202	-4.762	-5.653	-3,630	-3.692		-3.317	-3.916	-2.469	-2.417	1.634	-0.59	0.295	1.041	0.0	200	-1.459	1.162	1.500	1.482	1.896	2.470	0.380	3.654	7.50	6.063	6.890	6.617	9.00	4.0.4 4.0.4	8.015	9.808	9.5	7.036	7.179	6.453	5.012	4.083	4.646	7.75	2, 400	£.637	8.99	9.511	9.235	10,714	10.200
	1.738	7.710	23	Ş	200			92	699	173	986	3	2 2	3 5	15	3.46	412	23	7 6 9	3	5	907	2 3	?	921	953	3	334		8	?	300	3,4	200	151	٠,		7		***	٦.		: -:	•	_		•: •	-, -	. •	:		77	-16.696	-	-16.382
•	2	2	23	ž	25		: 3	539	980	633	519	000	92	200	23	503	639	368	200	606	695	182	668	::	131	433	310	203	833	370	603	358	120	922	105	23			5	090	99	70	5.5	723	3	5	18.223		36.	566	996	100	16,688	12,217	17.509
•	250	250	250	2 50	250	721	25.	251	251	251	251	252	252	25.2	252	252	253	253	253	254	254	256	254	72.	254	254	254	255	255	255	255	255	755 956	226	256	256	256	256	256	256	256	253	257	757	757	257	£ :	227	25.7	25.7	257	258	258	258	25 8
	3	31.	311	=======================================	: :	147	VAL	VAL	VAL	WE	7	SER	SER	200	SER	SER	GLY	CLY	֡֝֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓	176	LYS	LYS	176	27	LYS	LYS		¥ :	X X	Z Z	WAL	יי האר) (2 2	ARG	ARG.	ARG.	7 60	ARG	ARG	2	3 6	בונם	PHE	끝	<u> </u>	¥ :	31.	2 2	2	٠.	7	ž (780 680	2 0
		50						. :	2			2				0		ಶ	ט כ) z	ర	£	8	3 6	2	ပ	0	2 6	5 5	8	ĝ	U (0 2	. ე	8	ဗ	8	ב ב	Ę	23	، ن	2	. ວົ	8	8	ē	3	5 8	֓֞֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓	ن د	0	×	පි ද	5 E	98
									1957			1960		7067		•	•	•	0 Y	9 -	•	~							1980																8	_	3 3	3 8	3 8	Ö	8	0		2020	2022
	NTOK	20	10.5	N 04.	ATO.			¥ 5	ATOK	A TOK	ATOH	ATOH	A10H	A104	ATOK	ATOX	ž	¥		A TOK	ATOX	ATOK	ATOH	1014	ATON MOTO	ATOH	ATOH	A104	A TOX	ATOK	ATON	ATOH		7. E	ATON	ATOK	ADCA ,	- •	A 1 GH	ATOH	A10H		ATOX	A TOX	ATOK	ATOX	ATOK	101	300	A10X	ATOK	ATOK	ATOK	HOL A	7.10H



```
ATTOCK 2223 CA 1112 223 6.645 -14.575 -4.455 1.00 17.40 ATTOCK 2225 CA 1112 223 6.645 -14.575 -4.625 1.00 17.40 ATTOCK 2225 CA 1112 223 6.645 -14.575 -4.625 1.00 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40 17.40
```





```
ATTOCK 2533 CA PRO 222 -5.299 4.264 16.330 1.00 17.60 ANY 2535 CA PRO 222 -5.293 4.264 16.330 1.00 17.60 ANY 2535 CA PRO 222 -5.293 4.264 16.330 1.00 17.60 ANY 2535 CA PRO 222 -5.293 1.262 11.270 17.60 17.60 ANY 2535 CA CAY 252 CA
```

MARK 2466 CB PRO 313 -3.444 -5.421 9.423 1.00 17.6

MARK 2715 C PRO 313 -3.244 -7.531 14.61 1.00 17.6

MARK 2717 C PRO 313 -3.246 -7.531 14.61 1.00 17.6

MARK 2717 C PRE 314 -2.010 -2.111 10.111 1.10 17.6

MARK 2717 C PRE 314 -2.010 -2.111 10.111 1.10 17.6

MARK 2717 C PRE 314 -2.010 -2.111 10.111 1.10 17.6

MARK 2717 C PRE 314 -2.010 -2.111 10.111 1.10 17.6

MARK 2717 C PRE 314 -2.010 -2.111 10.111 1.10 17.6

MARK 2717 C PRE 314 -2.010 -2.111 10.111 1.10 17.6

MARK 2717 C PRE 314 -2.010 -2.111 10.111 1.10 17.6

MARK 2717 C PRE 314 -2.010 -2.111 10.111 1.10 17.6

MARK 2717 C PRE 314 -2.010 -2.111 10.11 1.10 17.6

MARK 2717 C PRE 314 -2.010 -2.111 10.11 1.10 17.6

MARK 2718 C PRE 314 -2.010 -2.111 17.11 17

```
ATOM. 2777 GO THE 332 24.277 -12.575 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00 11.00
```

```
ATOM 2922 CA NIE 369 24.349 6.687 -6.795 1.00 17.60 ATOM 2923 CB NIS 369 23.865 6.052 -7.746 1.00 17.60 ATOM 2923 CB NIS 369 21.852 -7.746 1.00 17.60 ATOM 2925 CD NIE 369 21.853 7.571 -9.397 1.00 17.60 ATOM 2925 CEL NIE 369 24.063 7.867 -9.397 1.00 17.60 ATOM 2927 CEL NIE 369 24.063 7.867 -9.397 1.00 17.60 ATOM 2927 CEL NIE 369 24.057 8.187 -9.397 1.00 17.60 ATOM 2929 C NIE 369 22.793 8.217 -10.414 1.00 17.60 ATOM 2921 C NIE 369 22.793 8.217 -10.414 1.00 17.60 ATOM 2921 C NIE 369 22.930 8.697 -6.651 1.00 17.60 ATOM 2931 CB AFP 370 25.097 8.627 -6.651 1.00 17.60 ATOM 2931 CC AFP 370 25.602 11.006 -6.861 1.00 17.60 ATOM 2931 CC AFP 370 25.602 11.006 -6.861 1.00 17.60 ATOM 2931 C AFP 370 25.404 10.619 -5.222 1.00 17.60 ATOM 2931 C AFP 370 25.404 10.619 -5.222 1.00 17.60 ATOM 2931 C AFP 370 25.798 10.619 -6.815 1.00 17.60 ATOM 2931 C AFP 370 25.798 10.619 -6.815 1.00 17.60 ATOM 2931 C AFP 370 25.798 10.20 27.20 1.00 17.60 ATOM 2931 C AFP 370 25.798 10.20 27.20 1.00 17.60 ATOM 2931 C AFP 370 26.399 9.719 -0.502 1.00 17.60 ATOM 2931 C AFP 370 26.399 9.719 -0.502 1.00 17.60 ATOM 2931 C AFP 370 26.399 9.719 -0.502 1.00 17.60 ATOM 2931 C AFP 370 26.399 9.719 -0.502 1.00 17.60 ATOM 2931 C AFP 370 26.399 9.719 -0.502 1.00 17.60 ATOM 2931 C AFP 370 26.399 9.719 -0.502 1.00 17.60 ATOM 2931 C AFP 370 26.399 9.719 -0.502 1.00 17.60 ATOM 2931 C AFP 370 26.399 9.719 -0.502 1.00 17.60 ATOM 2931 C AFP 370 26.399 9.719 -0.502 1.00 17.60 ATOM 2931 C AFP 370 26.399 9.719 -0.502 1.00 17.60 ATOM 2931 C AFP 370 26.399 9.719 -0.502 1.00 17.60 ATOM 2931 C AFP 370 26.390 9.719 -0.502 1.00 17.60 ATOM 2931 C AFP 370 26.390 9.719 -0.502 1.00 17.60 ATOM 2931 C AFP 370 26.390 9.719 -0.502 1.00 17.60 ATOM 2931 C AFP 370 27.781 9.719 -0.502 1.00 17.60 ATOM 2931 C AFP 370 27.781 9.719 -0.502 1.00 17.60 ATOM 2931 C AFP 370 27.781 9.719 -0.502 1.00 17.60 ATOM 2931 C AFP 370 27.781 9.719 -0.502 1.00 17.60 ATOM 2931 C AFP 370 27.781 9.710 17.60 ATOM 2931 C AFP 370 27.781 9.710 17.60 ATOM 2931 C AFP 370 27.781 9.710 17.60 ATOM 2931 C AFP 370 27.781 9.
```

```
| The color of the
```

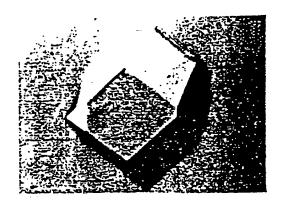
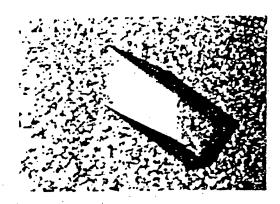


FIGURE 18A



FIGURE 18B



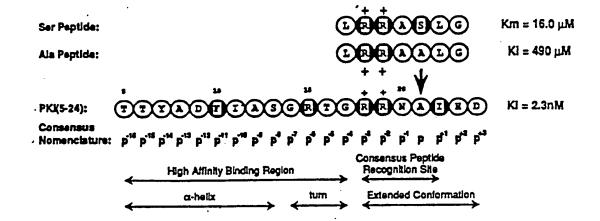
FIGURE

18C

Structure Solution Statistics

TABLE

Rsym	0.061	0.040	0.063	0.075	0.048				46/49								Data Selection	10-2.7 Å, F/6 > 2	$10-2.7 \text{ Å, } F/\sigma > 2$	$10-2.7 \text{ Å, F/}\sigma > 2$	40-27 A
cncss							176	3.01	0.43	2.05	0.60	0.81		1.51	0.61			=	=	ual	ual
Completeness (%)	98.1	87.3	76.1	92.1	91.1		0	3.8/	0.47	2.26	0.78	0.70		1.83	0.65		Ø	overall	overall	individual	individual
<[/0	2.9	13.2	13.3	6.9	11.1	•	(4) 130	4.20	0.53	2.56	0.89	0.64		2.08	0.71		tor				
2/1>	ä	~	-	•	-		Resolution (Å)	4.03	0.55	2.54	1.05	0.78		2.34	0.75		Final R-factor	0,304	0.228	0.195	0.212
No. of Reflections	12713	11291	7233	8809	11840		Shell	2775	0.62	3.00	1.19	0.51		2.81	0.64		Fina		-		
	:	_	,	x	-		Average	0.11	9.0	3.70	1.41	0.45		3.35	0.58		t-factor	73	74	ŧ	12
No. of Measurements	58889	27067	30973	23476	26464		7	0/./	0.75	3.80	1.68	0.38		3.72	0.53		Initial R-factor	0.473	0.434		0.221
								11.72	0.74	3.27	1.53	0.34		3.89	0.37		of s/Chains				•
dmin (Å)	2.7	2.7	3.0	3.0	2.7			Overall	0.57	2.73	96.0	0.50		2.26	09.0		No. of Residues/Cl	275/4	356/2	356/2	356/2
No. of Crystals	7				-	tatistics:			merit	fh/Eiso	Eanom			fh/Eiso				ed partial	ed full		
Diffraction Data: Data Sets	Native-1	Native-2	PHMB-1	PHMB-2	MgATP	SIRAS Phasing Statistics:	•		Mean figure of merit PIIMB-1	acentric r.m.s. fh/Eiso	r.m.s. AFanom/Eanom	Rc	PHMB-2	acentric r.m.s. fh/Eiso	$R_{\mathbf{c}}$	Refinement:	Model	A. First unrefined	B. First unrefined	C Latest X-PLOR	D. TNT



TABLE

PCT/US92/06137

48/49

TABLE 3

POINTS OF CONTACT	POSITION	CAPK	<u>CKII</u>
P+1	197 198 199 200 201 202 203 204 205	Thr Leu Cys Gly Thr Pro Glu Tyr Leu	Val Arg Val Ala Ser Arg Tyr Phe Lys
P-2	170 230	Glu Glu	His Glu
P-3	127 331	Glu Glu	Asp
P-6	203	Leu	
P-11	235 236 237 238 239	Tyr Pro Pro Phe Phe	

49/49

TABLE 4

Angstroms apart	Atom 1	
5.29	ASP 184 CA	GLY 186 CA
5.73	GLU 91 CA	GLY 186 CA
6.46	ASN 171 CA	ASP 184 CA
7.41	ASN 171 CA	ASP 166 CA
7.61	ASP 166 CA	GLY 186 CA
7.87	ASP 184 CA	GLU 91 CA
8.20	ASP 166 CA	ASP 184 CA
9.20	ASP 184 CA	LYS 72 CA
9.90	GLY 52 CA	LYS 72 CA
10.15	ASN 171 CA	GLY 186 CA
10.29	ASP 184 CA	GLY 52 CA
10.53	GLY 52 CA	GLY 186 CA
10.78	ASN 171 CA	GLY 52 CA
10.91	GLY 186 CA	LYS 72 CA
11.29	GLU 91 CA	LYS 72 CA
11.80	ARG 280 CA	GLU 208 CA
12.27	ASP 166 CA	GLU 91 CA
12.65	ASP 166 CA	GLY 52 CA
13.52	ASN 171 CA	LYS 72 CA
14.07	ASN 171 CA	GLU 91 CA
15.02	GLU 91 CA	GLY 52 CA
15.07	ASP 166 CA	GLU 208 CA
16.54	ASP 165 CA	LYS 72 CA
18.58	ARG 280 CA	ASP 166 CA
19.99	GLU 208 CA	GLY 186 CA
22.00	ASN 171 CA	GLU 208 CA
22.82	ASP 184 CA	GLU 208 CA
23.37	GLU 91 CA	GLU 208 CA
23.49	ARG 280 CA	ASN 171 CA
24.87	ARG 280 CA	GLY 186 CA
25.18	GLU 208 CA	GLY 52 CA
25.61	ARG 280 CA	ASP 184 CA
27.34	ARG 280 CA	GLU 91 CA
30.53	GLU 208 CA	LYS 72 CA
30.83	ARG 280 CA	GLY 52 CA
34.67	ARG 280 CA	LYS 72 CA

ç

T. 62 + 600			International Application No	<u> </u>
		ECT MATTER (if several classification		
	. 5 C12Q1/00 C07K13/0		C12N9/99; (G01N33/68
II. FIELDS	SEARCHED			
Cl. id. i		Minimum Do	ocumentation Searched?	
Classificati	ion System	·	Classification Symbols	
Int.Cl.	. 5	C12Q ; C12N ; C07K	G01N ; A61K	
			other than Minimum Documentation ents are Included in the Fields Searched ⁸	
		D TO BE RELEVANT ⁹		
Category °	Citation of Do	cument, ¹¹ with indication, where appr	ropriate, of the relevant passages 12	Relevant to Claim No.13
\ -	24 Augus	907 654 (PROGENICS PH st 1989 ms; example 1	HARMACEUTICALS)	1
		59 981 (BOEHRINGER) 1990		1,29
			-/~-	
	* - 10		,	
			•	j.
1				
1				
l				
-				
25		. 10		
•	ategories of cited docu nent defining the gene	ments: " rai state of the art which is not	"I" later document published after the interne or priority date and not in conflict with t	he application but
coasi	dered to be of particul:	ar relevance ned on or after the international	cited to understand the principle or theor invention	
uning	date	doubts on priority claim(s) or	"X" document of particular relevance; the cial cannot be considered novel or cannot be involve an inventive step	med invention considered to
which	is cited to establish the n or other special reas	e publication date of another	"Y" document of particular relevance; the clai	med invention
"O" docum		al disclosure, use, exhibition or	cannot be considered to involve an invent document is combined with one or more o ments, such combination being obvious to	ther such docu-
"P" docum		the international filing date but daimed	"A" document member of the same patent fam	· ·
. CERTIFIC		International Co		
HE OF THE AC	tual Completion of the		Date of Mailing of this International Sear	ch Report
	04 NOVEMBE	ΤΆΆζ	1 6. 11. 92	
ernational S	earching Authority		Signature of Authorized Officer	
	EUROPEAN	PATENT OFFICE	DELANGHE L.L.M.	

III. DOCUMI	ENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)	
Category °	Citation of Document, with Indication, where appropriate, of the relevant passages	Relevant to Claim No.
Å	CHEMICAL ABSTRACTS, vol. 105, no. 21, 24 November 1986, Columbus, Ohio, US; abstract no. 186487d, CLORE, G.MARIUS ET AL. 'Stereochemistry of binding of the tetrapeptide acetyl-Pro-Ala-Pro-Tyr-NH2 to porcine pancreatic elastase.Combined use of two-dimensional transferred nuclear Overhauser enhancement measurements, restrained molecular dynamics, X-ray crystallography and molecular modelling.' page 314; see abstract & J.MOL.BIOL.	Relevant to Claim No.
P,X	vol. 190, no. 2, 1986, ENG pages 259 - 267 CHEMICAL ABSTRACTS, vol. 115, no. 13, 30 September 1991, Columbus, Ohio, US; abstract no. 130637s, KNIGHTON, DANIEL R. ET AL. 'Crystallization of cAMP-dependent protein kinase.Cocrystals of the catalytic subunit with a 20 amino acid residue peptide inhibitor and magnesium-ATP diffract to 3.0 A resolution.' page 453; see abstract & J.MOL.BIOL. vol. 220, no. 2, 1991, ENG pages 217 - 220	1
P,X	SCIENCE vol. 253, no. 5018, 26 July 1991, LANCASTER, PA US pages 414 - 420 D.R.KNIGHTON ET AL. 'Structure of a peptide inhibitor bound to the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase.' see page 420 see the whole document	1-66

ANNEX TO THE INTERNATIONAL SEARCH REPORT ON INTERNATIONAL PATENT APPLICATION NO. US 9206137 SA 62983

This amex lists the patent family members relating to the patent documents cited in the above-mentioned international search report. The members are as contained in the European Patent Office EDP file on

The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information. 04/11/92

Patent document cited in search report	Publication date	1	Patent family member(s)	Publication date
WO-A-8907654	24-08-89	US-A- AU-B- AU-A- EP-A- JP-T-	4980281 612948 3184089 0403506 3503598	25-12-90 18-07-91 06-09-89 27-12-90 15-08-91
EP-A-0359981	28-03-90	DE-A- AU-B- AU-A- JP-A-	3827974 627207 3955089 2108636	22-02-90 20-08-92 22-02-90 20-04-90